PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/40, 15/82, 1/21, 15/10 — A1

(11) International Publication Number: **WO 96/21019**

(43) International Publication Date: 11 July 1996 (11.07.96)

(21) International Application Number: PCT/US95/07272

(22) International Filing Date: 7 June 1995 (07.06.95)

(30) Priority Data:
08/366,881 30 December 1994 (30.12.94) US

(60) Parent Application or Grant
(63) Related by Continuation
US 08/366,881 (CIP)
Filed on 30 December 1994 (30.12.94)

(71) Applicants *(for all designated States except US)*: ASGROW SEED COMPANY [US/US]; 2605 East Kilgore Road, Kalamazoo, MI 49002 (US). CORNELL RESEARCH FOUNDATION, INC. [US/US]; 20 Thornwood Drive, Ithaca, NY 14850 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: McMASTER, J., Russell [US/US]; 9432 Main Street, Galesburg, MI 49053 (US). BOESHORE, Maury, L. [US/US]; 8901 North 24th Street, Kalamazoo, MI 49004 (US). TRICOLI, David, M. [US/US]; 2332 S. Rose Street, Kalamazoo, MI 49001 (US). REYNOLDS, John, F. [US/US]; 14815 Trillium Drive, Augusta, MI 49012 (US). CARNEY, Kim, J. [US/US]; 8607 East B Avenue, Richland, MI 49083 (US). GONSALVES, Dennis [US/US]; 595 Castle Street, Geneva, NY 14456 (US).

(74) Agent: PERRY, Lawrence, S.; Fitzpatrick, Cella, Harper & Scinto, 277 Park Avenue, New York, NY 10172 (US).

(81) Designated States: AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LU, LV, MD, MG, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SI, SK, TJ, TT, UA, US, UZ, VN, ARIPO patent (KE, MW, SD, SZ, UG), European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: PAPAYA RINGSPOT VIRUS COAT PROTEIN GENE

(57) Abstract

A coat protein gene of papaya ringspot virus strain FLA83 W is provided.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	**GB**	United Kingdom	**MW**	Malawi
AT	Austria	**GE**	Georgia	**MX**	Mexico
AU	Australia	**GN**	Guinea	**NE**	Niger
BB	Barbados	**GR**	Greece	**NL**	Netherlands
BE	Belgium	**HU**	Hungary	**NO**	Norway
BF	Burkina Faso	**IE**	Ireland	**NZ**	New Zealand
BG	Bulgaria	**IT**	Italy	**PL**	Poland
BJ	Benin	**JP**	Japan	**PT**	Portugal
BR	Brazil	**KE**	Kenya	**RO**	Romania
BY	Belarus	**KG**	Kyrgystan	**RU**	Russian Federation
CA	Canada	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CF	Central African Republic			**SE**	Sweden
CG	Congo	**KR**	Republic of Korea	**SG**	Singapore
CH	Switzerland	**KZ**	Kazakhstan	**SI**	Slovenia
CI	Côte d'Ivoire	**LI**	Liechtenstein	**SK**	Slovakia
CM	Cameroon	**LK**	Sri Lanka	**SN**	Senegal
CN	China	**LR**	Liberia	**SZ**	Swaziland
CS	Czechoslovakia	**LT**	Lithuania	**TD**	Chad
CZ	Czech Republic	**LU**	Luxembourg	**TG**	Togo
DE	Germany	**LV**	Latvia	**TJ**	Tajikistan
DK	Denmark	**MC**	Monaco	**TT**	Trinidad and Tobago
EE	Estonia	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**UG**	Uganda
FI	Finland	**ML**	Mali	**US**	United States of America
FR	France	**MN**	Mongolia	**UZ**	Uzbekistan
GA	Gabon	**MR**	Mauritania	**VN**	Viet Nam

TITLE

PAPAYA RINGSPOT VIRUS COAT PROTEIN GENE

Field of the Invention

This invention relates to a coat protein gene derived from papaya ringspot virus. More specifically, the invention relates to the genetic engineering of plants and to a method for conferring viral resistance to a plant using an expression cassette encoding papaya ringspot virus PRV FLA83 W coat protein.

Background of the Invention

Many agriculturally important crops are susceptible to infection by plant viruses, particularly papaya ringspot virus, which can seriously damage a crop, reduce its economic value to the grower, and increase its cost to the consumer. Attempts to control or prevent infection of a crop by a plant virus such as papaya ringspot virus have been made, yet viral pathogens continue to be a significant problem in agriculture.

Scientists have recently developed means to produce virus resistant plants using genetic engineering techniques. Such an approach is advantageous in that the genetic material which provides the protection is incorporated into the genome of the plant itself and can be passed on to its progeny. A host plant is resistant if it possesses the ability to suppress or retard the multiplication of a virus, or the development of pathogenic symptoms. "Resistant" is the opposite of "susceptible," and may be divided into: (1) high, (2) moderate, or (3) low resistance, depending upon its effectiveness. Essentially, a resistant plant shows reduced or no symptom expression, and virus multiplication within it is reduced or negligible. Several different types of host resistance to viruses are recognized. The host may be resistant to: (1) establishment of infection, (2) virus multiplication, or (3) viral movement.

Potyviruses are a distinct group of plant viruses which are pathogenic to various crops, and which demonstrate cross-infectivity between plant members of different families. Generally, a potyvirus is a single-stranded RNA virus that is surrounded by a repeating protein monomer, which is termed the coat protein (CP). The majority of the potyviruses are transmitted in a nonpersistent manner by aphids. As can be seen from the wide range of crops affected by potyviruses, the host range includes such diverse families of plants as Solanaceae, Chenopodiaceae, Gramineae, Compositae, Leguminosae, Dioscroeaceae, Cucurbitaceae, and Caricaceae. Potyviruses include watermelon mosaic virus II (WMVII); zucchini yellow mosaic virus (ZYMV), potato virus Y, tobacco etch and many others.

Another potyvirus of economic significance is papaya ringspot virus (PRV). Two groups of PRV have been

identified: the "P" or "papaya ringspot" type infects papayas; and the "W" or "watermelon" type infects cucurbits, e.g., squash, but it is unable to infect papaya. Thus, these two groups can be distinguished by host range differences.

The potyviruses consist of flexous, filamentous particles of dimensions approximately 780 x 12 nanometers. The viral particles contain a single-stranded positive polarity RNA genome containing about 10,000 nucleotides. Translation of the RNA genome of potyviruses shows that the RNA encodes a single large polyprotein of about 330 kD. This polyprotein contains several proteins; these include the coat protein, nuclear inclusion proteins NIa and NIb, cytoplasmic inclusion protein (CI), and other proteases and movement proteins. These proteins are found in the infected plant cell and form the necessary components for viral replication. One of the proteins contained in the polyprotein is a 35 kD capsid or coat protein which coats and protects the viral RNA from degradation. One of the nuclear inclusion proteins, NIb, is an RNA replicase component and is thought to have polymerase activity. CI, a second inclusion protein, is believed to participate in the replicase complex and have a helicase activity. NIa, a third inclusion protein, has a protease activity. In the course of potyvirus infection, NIa and NIb are translationally transported across the nuclear membrane into the nucleus of the infected plant cell at the later stages of infection and accumulate to high levels.

The location of the protease gene appears to be conserved in these viruses. In the tobacco etch virus, the protease cleavage site has been determined to be the dipeptide Gln-Ser, Gln-Gly, or Gln-Ala.

Conservation of these dipeptides at the cleavage sites in these viral polyproteins is apparent from the sequences of the above-listed potyviruses.

Expression of the coat protein genes from tobacco mosaic virus, alfalfa mosaic virus, cucumber mosaic virus, and potato virus X, among others, in transgenic plants has resulted in plants which are resistant to infection by the respective virus. For reviews, see Fitchen et al., Annu. Rev. Microbiol., 47, 739 (1993) and Wilson, Proc. Natl. Acad. Sci. USA, 90, 3134 (1993). For papaya ringspot virus, Ling et al. (Bio/Technology, 9, 752 (1991)) found that transgenic tobacco plants expressing the PRV coat protein gene isolated from the PRV strain HA 5-1 (mild) showed delayed symptom development and attenuation of symptoms after infection by a number of potyviruses, including tobacco etch (TEV), potato virus Y (PVY), and pepper mottle virus (PeMV). PRV does not infect tobacco, however. Thus, PRV CP transgenic tobacco plants cannot be used to evaluate protection against PRV. Fitch et al. (Bio/Technology, 10, 1466 (1992)), Gonsalves (American J. of Bot., 79, 88 (1992)), and Lius et al., 91st Annual Meeting of the American Society for Horticultural Science Hortscience, 29, 483 (1994)) reported that four R_0 papaya plants made transgenic for a PRV coat protein gene taken from strain HA 5-1 (mild) displayed varying degrees of resistance against PRV infection, and one line (S55-1) appeared completely resistant to PRV. This appears to be the only papaya line that shows complete resistance to PRV infection.

Thus, there is a continuing need for the transgenic expression of genes derived from potyviruses at levels which confer resistance to infection by these viruses.

SUMMARY OF THE INVENTION

This invention provides an isolated and purified DNA molecule that encodes the coat protein for the FLA83 W-type strain of papaya ringspot virus (PRV). The invention also provides a chimeric expression cassette comprising this DNA molecule, a promoter which functions in plant cells to cause the production of an RNA molecule, and at least one polyadenylation signal comprising 3' nontranslated DNA which functions in plant cells to cause the termination of transcription and the addition of polyadenylated ribonucleotides to the 3' end of the transcribed mRNA sequences, wherein the promoter is operably linked to the DNA molecule, and the DNA molecule is operably linked to the polyadenylation signal. Another embodiment of the invention is exemplified by the insertion of multiple virus gene expression cassettes into one purified DNA molecule, e.g., a plasmid. Preferably, these cassettes include the promoter of the 35S gene of cauliflower mosaic virus and the polyadenylation signal of the cauliflower mosaic virus 35S gene.

Also provided are bacterial cells, and transformed plant cells, containing the chimeric expression cassettes comprising the coat protein gene derived from the FLA83 W-type strain of papaya ringspot virus (referred to herein as PRV FLA83 W), and preferably the 35S promoter of cauliflower mosaic virus and the polyadenylation signal of the cauliflower mosaic virus 35S gene. Plants are also provided, wherein the plants comprise a plurality of transformed cells transformed with a cassette containing the coat protein gene derived from the PRV FLA83 W strain, and preferably the cauliflower mosaic virus 35S promoter and the polyadenylation signal of the cauliflower mosaic virus gene. Transformed plants of this invention

include tobacco, corn, cucumber, peppers, potatoes, soybean, squash, and tomatoes. Especially preferred are members of the *Cucurbitaceae* (e.g., squash and cucumber) family.

Another aspect of the present invention is a method of preparing a PRV-resistant plant, such as a dicot, comprising: transforming plant cells with a chimeric expression cassette comprising a promoter functional in plant cells operably linked to a DNA molecule that encodes a coat protein as described above; regenerating the plant cells to provide a differentiated plant; and identifying a transformed plant that expresses the PRV coat protein at a level sufficient to render the plant resistant to infection by the specific strains of PRV disclosed herein.

As used herein, with respect to a DNA molecule or "gene," the phrase "isolated and purified" is defined to mean that the molecule is either extracted from its context in the viral genome by chemical means and purified and/or modified to the extent that it can be introduced into the present vectors in the appropriate orientation, i.e., sense or antisense. As used herein, the term "chimeric" refers to the linkage of two or more DNA molecules which are derived from different sources, strains or species (e.g., from bacteria and plants), or the linkage of two or more DNA molecules, which are derived from the same species and which are linked in a way that does not occur in the native genome. As used herein, the term "expression" is defined to mean transcription or transcription followed by translation of a particular DNA molecule.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1. The nucleotide sequence of the coat protein gene (long version) of PRV FLA83 W [SEQ ID NO:1]. The amino acid sequence of the encoded open reading frame is shown below the nucleotide sequence [SEQ ID NO:2].

Fig. 2. The nucleotide sequence of the coat protein gene (short version) of PRV FLA83 W [SEQ ID NO:3]. The amino acid sequence of the encoded open reading frame is shown below the nucleotide sequence [SEQ ID NO:4].

Fig. 3. The alignment of the nucleotide sequences of the PRV FLA83 W long (LG) and short (SH) coat protein genes [SEQ ID NOS:1 and 3]. The primer pairs RMM384-385 and RMM388-385 are shown [SEQ ID NOS:5, 6, and 7]. The primer pairs RMM384-385 and RMM388-385 were used to PCR amplify and install novel NcoI restriction sites for LG and SH coat protein genes, respectively. The viral-specific sequences present in RMM384, RMM385, and RMM388 are homologous to sequences in PRV HA (attenuated) USA P (Quemada et al., J. Gen. Virol., 71, 203 (1990)). In addition, all three oligomers contain novel NcoI sites (underlined sequences).

Fig. 4. The alignment of the coat protein coding sequences from papaya ringspot virus isolates: Australian W (Bateson et al., Arch-Viol, 123, 101 (1992)) [SEQ ID NO:8]; HA P (Yeh et al., J. Gen. Virol., 73:2531 (1992)) [SEQ ID NO:9]; USA P (Quemada et al., J. Gen. Virol., 71, 203 (1990)) [SEQ ID NO:10]; USA-W (Quemada et al., J. Gen. Virol., 71, 203 (1990)) [SEQ ID NO:11]; and FLA83 W SH [SEQ ID NO:3]. Alignments were generated using the UWGCG Pileup program. The dots represent either the lack of sequence information at the ends of the coat protein

gene or gaps in homology in sequences relative to others in the alignment.

Fig. 5. The alignment of the coat protein amino acid sequences from papaya ringspot virus isolates: Australian W [SEQ ID NO:12]; HA P [SEQ ID NO:13]; USA P [SEQ ID NO:14]; USA W [SEQ ID NO:15]; and FLA83 W LG and SH [SEQ ID NOS:2 and 4]. Alignments were generated using the UWGCG Pileup program. The dots represent either the lack of sequence information at the 5' end of the coat protein gene or gaps in homology in sequences relative to others in the alignment. Sequence homology differences between virus strains are underlined. The deduced amino acid sequence of the PRV FLA83 W coat protein (CP) gene disclosed here is unique compared with the PRV coat protein amino acid sequences of the four strains shown in the figure. The PRV FLA83 CP amino acid sequence differs from all other published PRV CP sequences in at least 14 positions (Numbers 1-14). The FLA83 W CP gene possesses a 6-bp insertion (Figure 4) relative to other PRV CP genes characterized to date (see "*INSERTION*" in Figure 5). This 6-bp insertion codes for the amino acids threonine-threonine.

Fig. 6. A schematic diagram of the cloning strategy for the long version of PRV coat protein gene (PRVFLA83cp16[s] and [as]). Single stranded cDNA was produced with PRV virion RNA as template and reverse transcriptase. After PCR amplification, the PCR product was digested with NcoI and inserted into the NcoI site of pUC18cpexpress to yield sense and antisense constructs.

Fig. 7. A schematic representation of the cloning strategy for the short version of PRV coat protein (PRVFLA83cp34 [s] and [as]). Single stranded cDNA was

produced with PRV virion RNA as template and reverse transcriptase. After PCR amplification, the PCR product was digested with NcoI and inserted into the NcoI site of pUC18cpexpress to yield sense and antisense constructs.

Fig. 8. The alignment of nucleotide sequences for seven isolates of PRV.

Fig. 9. The alignment of amino acid sequences for seven isolates of PRV.

Fig. 10. The theoretical relations between the seven PRV isolates of Figures 8 and 9.

DETAILED DESCRIPTION OF THE INVENTION

Papaya Ringspot Virus (PRV) is a single-stranded (+) RNA plant virus. The viral genome is approximately 10,000 bases in length. The expression strategy of potyviruses includes translation of a complete polyprotein from the positive sense viral genomic RNA. Translation of the genomic RNA produces a 330 kD protein which is subsequently cleaved into at least seven smaller viral proteins by a virally encoded protease. The virally encoded proteins include a 35 kD protein at the amino terminal end of the 330 kD protein which is thought to be involved in cell to cell transmission, H C protein is 56 kD in size and is believed to be involved in insect transmission and possess proteolytic activity, a 50 kD protein, a 90 kD cylindrical inclusion protein (CI) which is part of the replicase complex and possesses helicase activity, a 6 kD VPg protein which is covalently attached to the 5' end of the viral genomic RNA, a 49 kD NIa protein which functions as a protease, a 60 kD NIb protein which

functions as a polymerase, and the coat protein (36 kD).

Two types of PRV have been established based on host range. One type is designated "P type"; it infects *Caricacae* (e.g., papaya), *Cucurbitaceae* (e.g., cucurbitis), and *Chenopodiaceae* (e.g., *Chenopodium*) (Wang et al., Phytopathology, 84, 1205 (1994)). A second type is designated "W type"; it infects only *Cucurbitaceae* and *Chenopodiaceae* (Wang et al., Phytopathology, 84, 1205 (1994)). Isolates of the P type include HA-severe (Wang et al., Virus Arch. Virol., 127, 345 (1992)), HA5-1, called USA P herein, YK (Wang et al., Phytopathology, 84, 1205 (1994)), and other isolates as described in Tennant et al. (Phytopathology, 84, 1359 (1994)). Isolates of the W type include FLA83, disclosed herein, PRV-W type (Yeh et al., Phytopath., 74, 1081 (1984)) and PRV-W (Aust) (Bateson et al., Arch-Viol, 123, 101 (1992)).

Previous work has shown that the potyvirus NIa protease cleaves the coat protein from the adjacent protein NIb (Restrepo-Hartwig et al., J. Virol., 66, 5662 (1992); Dougherty et al., Ann. Rev. Phytopath., 26, 123 (1988); Carrington et al., J. Virol., 61, 2540 (1987)). The determination of the N-terminal amino acid sequences of the coat protein have been problematic (Yeh et al., J. Gen. Virol., 73, 2531 (1992); Wang et al., Virus Arch. Virol., 127, 345 (1992)), therefore the amino terminus of the coat protein remains unclear. The sites predicted for the NIa/coat protein cleavage site are underlined in Figure 5 (VFHQ/SKNE in Quemada et al., J. Gen. Virol, 71, 203 (1990); VFHQ/SKNE in Bateson et al., Arch. Viol., 123, 101 (1992); VYHE/SRGTD in Yeh et al., J. Gen. Virol., 73, 2531 (1992); VLEQ/APFN and VFHQ/AKNE described herein).

To practice the present invention, the coat protein gene of a virus must be isolated from the viral genome and inserted into a vector. Thus, the present invention provides isolated and purified DNA molecules that encode the coat protein of PRV FLA83. As used herein, a DNA molecule that encodes a coat protein gene includes nucleotides of the coding strand, also referred to as the "sense" strand, as well as nucleotides of the noncoding strand, complementary strand, also referred to as the "antisense" strand, either alone or in their base-paired configuration. Thus, a DNA molecule that encodes the coat protein of PRV FLA83, for example, includes the DNA molecule having the nucleotide sequence of Figure 1 [SEQ ID NO:1], a DNA molecule complementary to the nucleotide sequence of Figure 1 [SEQ ID NO:1], as well as a DNA molecule which also encodes a PRV coat protein and its complement which hybridizes with a PRV FLA83-specific DNA probe in hybridization buffer with 6XSSC, 5X Denhardt's reagent, 0.5% SDS and 100 μg/mL denatured, fragmented salmon sperm DNA and remains bound when washed at 68°C in 0.1XSSC and 0.5% SDS (Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd ed. (1989)). Moreover, the DNA molecules of the present invention can include non-PRV coat protein nucleotides that do not interfere with expression. Preferably, the isolated and purified DNA molecules of the present invention comprise a single coding region for the coat protein. Thus, preferably the DNA molecules of the present invention are those "consisting essentially of" DNA that encodes the coat protein.

The PRV coat protein gene does not contain the signals necessary for its expression once transferred and integrated into a plant genome. Accordingly, a vector must be constructed to provide the regulatory sequences such that they will be functional upon inserting a

desired gene. When the expression vector/insert construct is assembled, it is used to transform plant cells which are then used to regenerate plants. These transgenic plants carry the viral gene in the expression vector/insert construct. The gene is expressed in the plant and increased resistance to viral infection is conferred thereby.

Several different methods exist to isolate a viral gene. To do so, one having ordinary skill in the art can use information about the genomic organization of potyviruses to locate and isolate the coat protein gene. The coat protein gene is located at the 3' end of the RNA, just prior to a stretch of about 200-300 adenine nucleotide residues. Additionally, the information related to proteolytic cleavage sites is used to determine the N-terminus of the potyvirus coat protein gene. The protease recognition sites are conserved in the potyviruses and have been determined to be either the dipeptide Gln-Ser, Gln-Gly, or Gln-Ala. The nucleotide sequences which encode these dipeptides can be determined.

Using methods well known in the art, a quantity of virus is grown and harvested. The viral RNA is then separated and a viral gene isolated using a number of known procedures. A cDNA library is created using the viral RNA, by methods known to the art. The viral RNA is incubated with primers that hybridize to the viral RNA and reverse transcriptase, and a complementary DNA molecule is produced. A DNA complement of the complementary DNA molecule is produced and that sequence represents a DNA copy (cDNA) of the original viral RNA molecule. The DNA complement can be produced in a manner that results in a single double stranded cDNA or polymerase chain reactions can be used to amplify the DNA encoding the cDNA with the use of

oligomer primers specific for the coat protein. These primers can include novel restriction sites used in subsequent cloning steps. Thus, a double stranded DNA molecule is generated which contains the sequence information of the viral RNA. These DNA molecules can be cloned in *E. coli* plasmid vectors after the additions of restriction enzyme linker molecules by DNA ligase. The various fragments are inserted into cloning vectors, such as well-characterized plasmids, which are then used to transform *E. coli* and create a cDNA library.

Previously isolated PRV coat protein genes can be used as hybridization probes to screen the cDNA library to determine if any of the transformed bacteria contain DNA fragments with sequences coding for the PRV coat protein region. The cDNA inserts in any bacterial colonies which contain this region can be sequenced. The coat protein gene is present in its entirety in colonies which have sequences that extend 5' to a sequence which encodes a N-terminal proteolytic cleavage site and 3' to a stop codon.

Alternatively, cDNA fragments can be inserted in the sense orientation into expression vectors. Antibodies against the coat protein can be used to screen the cDNA expression library and the gene can be isolated from colonies which express the protein.

Another molecular strategy to provide virus resistance in transgenic plants is based on antisense RNA. As is well known, a cell manufactures protein by transcribing the DNA of the gene encoding that protein to produce RNA, which is then processed to messenger RNA (mRNA) (e.g., by the removal of introns) and finally translated by ribosomes into protein. This process may be inhibited in the cell by the presense of antisense

RNA. The term antisense RNA means an RNA sequence which is complementary to a sequence of bases in the mRNA in question in the sense that each base (or the majority of bases) in the antisense sequence (read in the 3' to 5' sense) is capable of pairing with the corresponding base (G with C, A with U) in the mRNA sequence read in the 5' to 3' sense. It is believed that this inhibition takes place by formation of a complex between the two complementary strands of RNA, thus preventing the formation of protein. How this works is uncertain: the complex may interfere with further transcription, processing, transport or translation, or degrade the mRNA, or have more than one of these effects. This antisense RNA may be produced in the cell by transformation of the cell with an appropriate DNA construct arranged to transcribe the non-template strand (as opposed to the template strand) of the relevant gene (or of a DNA sequence showing substantial homology therewith).

The use of antisense RNA to downregulate the expression of specific plant genes is well known. Reduction of gene expression has led to a change in the phenotype of the plant: either at the level of gross visible phenotypic difference, e.g., lack of anthocyanin production in flower petals of petunia leading to colorless instead of colored petals (van der Krol et al., Nature, 333:866-869 (1988)); or at a more subtle biochemical level, e.g., change in the amount of polygalacturonase and reduction in depolymerization of pectin during tomato fruit ripening (Smith et al., Nature, 334:724-726 (1988)).

Another more recently described method of inhibiting gene expression in transgenic plants is the use of sense RNA transcribed from an exogenous template to downregulate the expression of specific plant genes

(Jorgensen, Keystone Symposium "Improved Crop and Plant Products through Biotechnology," Abstract X1-022 (1994)). Thus, both antisense and sense RNA have been proven to be useful in achieving downregulation of gene expression in plants.

In the present invention, the DNA molecule encoding the coat protein gene of the papaya ringspot virus strain FLA83 has been determined and the gene has been inserted into an expression vector. These expression cassettes can be individually placed into a vector that can be transmitted into plants, preferably a binary vector. Alternatively, two or more PRV coat protein genes can each be present in an expression cassette which can be placed into the same binary vector, or a PRV coat protein expression cassette of the present invention can be placed into a binary vector with one or more viral gene expression cassettes. The expression vectors contain the necessary genetic regulatory sequences for expression of an inserted gene. The coat protein gene is inserted such that those regulatory sequences are functional and the genes can be expressed when incorporated into a plant genome. For example, vectors of the present invention can contain combinations of expression cassettes that include DNA from PRVcoat protein genes other than those of FLA83 (i.e., heterologous PRV coat protein genes), a cucumber mosaic virus coat protein gene, a squash mosaic virus coat protein gene, a zucchini yellow mosaic virus-2 coat protein gene, and a watermelon mosaic virus coat protein gene.

Moreover, when combinations of viral gene expression cassettes are placed in the same binary plasmid, and that multigene cassette containing plasmid transformed into a plant, the viral genes all preferably exhibit substantially the same degrees of efficacy when present

in transgenic plants. For example, if one examines numerous transgenic lines containing two different intact viral gene cassettes, the transgenic line will be immune to infection by both viruses. Similarly, if a line exhibits a delay in symptom development to one virus, it will also exhibit a delay in symptom development to the second virus. Finally, if a line is susceptible to one of the viruses it will be susceptible to the other. This phenomenon is unexpected. If there were not a correlation between the efficacy of each gene in these multiple gene constructs, this approach as a tool in plant breeding would probably be prohibitively difficult to use. Even with single gene constructs, one must test numerous transgenic plant lines to find one that displays the appropriate level of efficacy. The probability of finding a line with useful levels of expression can range from 10-50% (depending on the species involved). For further information refer to Applicants' Assignees copending Patent Application Serial No. 08/366,991 entitled "Transgenic Plants Expressing DNA Constructs Containing a Plurality of Genes to Impart Virus Resistance" filed on December 30, 1994, and incorporated by reference herein.

In order to express the viral gene, the necessary genetic regulatory sequences must be provided. Since the coat protein of a potyvirus is produced by the post-translational processing of a polyprotein, the coat protein gene isolated from viral RNA does not contain transcription and translation signals necessary for its expression once transferred and integrated into a plant genome. It must, therefore, be engineered to contain a plant expressible promoter, a translation initiation codon (ATG), and a plant functional poly(A) addition signal (AATAAA) 3' of its translation termination codon. In the present invention, the coat

proteins are inserted into vectors which contain cloning sites for insertion 3' of the initiation codon and 5' of the poly(A) signal. The promoter is 5' of the initiation codon such that when structural genes are inserted at the cloning site, a functional unit is formed in which the inserted genes are expressed under the control of the various genetic regulatory sequences.

The segment of DNA referred to as the promoter is responsible for the regulation of the transcription of DNA into mRNA. A number of promoters which function in plant cells are known in the art and can be employed in the practice of the present invention. These promoters can be obtained from a variety of sources such as plants or plant viruses, and can include, but are not limited to, promoters isolated from the caulimovirus group such as the cauliflower mosaic virus 35S promoter (CaMV35S), the enhanced cauliflower mosaic virus 35S promoter (enh CaMV35S), the figwort mosaic virus full-length transcript promoter (FMV35S), and the promoter isolated from the chlorophyll a/b binding protein. Other useful promoters include promoters which are capable of expressing the potyvirus proteins in an inducible manner or in a tissue-specific manner in certain cell types in which the infection is known to occur. For example, the inducible promoters from phenylalanine ammonia lyase, chalcone synthase, hydroxyproline rich glycoprotein, extensin, pathogenesis-related proteins (e.g. PR-1a), and wound-inducible protease inhibitor from potato may be useful.

Preferred promoters for use in the present viral gene expression cassettes include the constitutive promoters from CaMV, the Ti genes nopaline synthase (Bevan et al., Nucleic Acids Res. II, 369 (1983)) and octopine synthase (Depicker et al., J. Mol. Appl. Genet., 1, 561

(1982)), and the bean storage protein gene phaseolin. The poly(A) addition signals from these genes are also suitable for use in the present cassettes. The particular promoter selected is preferably capable of causing sufficient expression of the DNA coding sequences to which it is operably linked, to result in the production of amounts of the proteins effective to provide viral resistance, but not so much as to be detrimental to the cell in which they are expressed. The promoters selected should be capable of functioning in tissues including, but not limited to, epidermal, vascular, and mesophyll tissues. The actual choice of the promoter is not critical, as long as it has sufficient transcriptional activity to accomplish the expression of the preselected proteins and subsequent conferral of viral resistance to the plants.

The nontranslated leader sequence can be derived from any suitable source and can be specifically modified to increase the translation of the mRNA. The 5' nontranslated region can be obtained from the promoter selected to express the gene, an unrelated promoter, the native leader sequence of the gene or coding region to be expressed, viral RNAs, suitable eucaryotic genes, or a synthetic gene sequence. The present invention is not limited to the constructs presented in the following examples.

The termination region or 3' nontranslated region which is employed is one which will cause the termination of transcription and the addition of polyadenylated ribonucleotides to the 3' end of the transcribed mRNA sequence. The termination region can be native with the promoter region, native with the gene, or can be derived from another source, and preferably include a terminator and a sequence coding for polyadenylation. Suitable 3' nontranslated regions of the chimeric plant

gene include but are not limited to: (1) the 3' transcribed, nontranslated regions containing the polyadenylation signal of *Agrobacterium* tumor-inducing (Ti) plasmid genes, such as the nopaline synthase (NOS) gene; and (2) plant genes like the soybean 7S storage protein genes.

Preferably, the expression cassettes of the present invention are engineered to contain a constitutive promoter 5' to its translation initiation codon (ATG) and a poly(A) addition signal (AATAAA) 3' to its translation termination codon. Several promoters which function in plants are available, however, the preferred promoter is the 35S constitutive promoters from cauliflower mosaic virus (CaMV). The poly(A) signal can be obtained from the CaMV 35S gene or from any number of well characterized plant genes, i.e., nopaline synthase, octopine synthase, and the bean storage protein gene phaseolin. The constructions are similar to that used for the expression of the CMV C coat protein in PCT Patent Application PCT/US88/04321, published on June 29, 1989 as WO 89/05858, claiming the benefit of U.S.S.N. 135,591, filed December 21, 1987, entitled "Cucumber Mosaic Virus Coat Protein Gene," and the CMV WL coat protein in PCT Patent Application PCT/US89/03288, published on March 8, 1990 as WO 90/02185, claiming the benefit of U.S.S.N. 234,404, filed August 19, 1988, entitled "Cucumber Mosaic Virus Coat Protein Gene."

Selectable marker genes can be incorporated into the present expression cassettes and used to select for those cells or plants which have become transformed. The marker gene employed may express resistance to an antibiotic, such as kanamycin, gentamycin, G418, hygromycin, streptomycin, spectinomycin, tetracyline, chloramphenicol, and the like. Other markers could be

employed in addition to or in the alternative, such as, for example, a gene coding for herbicide tolerance such as tolerance to glyphosate, sulfonylurea, phosphinothricin, or bromoxynil. Additional means of selection could include resistance to methotrexate, heavy metals, complementation providing prototrophy to an auxotrophic host, and the like.

The particular marker employed will be one which will allow for the selection of transformed cells as opposed to those cells which are not transformed. Depending on the number of different host species, one or more markers can be employed, where different conditions of selection would be useful to select the different host, and would be known to those of skill in the art. A screenable marker such as the β-glucuronidase gene can be used in place of, or with, a selectable marker. Cells transformed with this gene can be identified by the production of a blue product on treatment with 5-bromo-4-chloro-3-indoyl-β-D-glucuronide (X-Gluc).

In developing the present expression construct, i.e., expression cassette, the various components of the expression construct such as the DNA molecules, linkers, or fragments thereof will normally be inserted into a convenient cloning vector, such as a plasmid or phage, which is capable of replication in a bacterial host, such as *E. coli*. Numerous cloning vectors exist that have been described in the literature. After each cloning, the cloning vector can be isolated and subjected to further manipulation, such as restriction, insertion of new fragments, ligation, deletion, resection, insertion, *in vitro* mutagenesis, addition of polylinker fragments, and the like, in order to provide a vector which will meet a particular need.

For *Agrobacterium*-mediated transformation, the expression cassette will be included in a vector, and flanked by fragments of the *Agrobacterium* Ti or Ri plasmid, representing the right and, optionally the left, borders of the Ti or Ri plasmid transferred DNA (T-DNA). This facilitates integration of the present chimeric DNA sequences into the genome of the host plant cell. This vector will also contain sequences that facilitate replication of the plasmid in *Agrobacterium* cells, as well as in *E. coli* cells.

All DNA manipulations are typically carried out in *E. coli* cells, and the final plasmid bearing the potyvirus gene expression cassette is moved into *Agrobacterium* cells by direct DNA transformation, conjugation, and the like. These *Agrobacterium* cells will contain a second plasmid, also derived from Ti or Ri plasmids. This second plasmid will carry all the viral genes required for transfer of the foreign DNA into plant cells. Suitable plant transformation cloning vectors include those derived from a Ti plasmid of *Agrobacterium tumefaciens*, as generally disclosed in Glassman et al. (U.S. Pat. No. 5,258,300), or *Agrobacterium rhizogenes*.

A variety of techniques are available for the introduction of the genetic material into or transformation of the plant cell host. However, the particular manner of introduction of the plant vector into the host is not critical to the practice of the present invention, and any method which provides for efficient transformation can be employed. In addition to transformation using plant transformation vectors derived from the tumor-inducing (Ti) or root-inducing (Ri) plasmids of *Agrobacterium*, alternative methods could be used to insert the DNA constructs of the present invention into plant cells. Such methods may

include, for example, the use of liposomes, transformation using viruses or pollen, chemicals that increase the direct uptake of DNA (Paszkowski et al., EMBO J., 3, 2717 (1984)), microinjection (Crossway et al., Mol. Gen. Genet., 202, 179 (1985)), electroporation (Fromm et al., Proc. Natl. Acad. Sci. USA, 82, 5824 (1985)), or high-velocity microprojectiles (Klein et al., Nature, 327, 70 (1987)).

The choice of plant tissue source or cultured plant cells for transformation will depend on the nature of the host plant and the transformation protocol. Useful tissue sources include callus, suspension culture cells, protoplasts, leaf segments, stem segments, tassels, pollen, embryos, hypocotyls, tuber segments, meristematic regions, and the like. The tissue source is regenerable, in that it will retain the ability to regenerate whole, fertile plants following transformation.

The transformation is carried out under conditions directed to the plant tissue of choice. The plant cells or tissue are exposed to the DNA carrying the present potyvirus multi-gene expression cassette for an effective period of time. This can range from a less-than-one-second pulse of electricity for electroporation, to a two-to-three day co-cultivation in the presence of plasmid-bearing Agrobacterium cells. Buffers and media used will also vary with the plant tissue source and transformation protocol. Many transformation protocols employ a feeder layer of suspended culture cells (tobacco or Black Mexican Sweet Corn, for example) on the surface of solid media plates, separated by a sterile filter paper disk from the plant cells or tissues being transformed.

Following treatment with DNA, the plant cells or tissue may be cultivated for varying lengths of time prior to selection, or may be immediately exposed to a selective agent such as those described hereinabove. Protocols involving exposure to *Agrobacterium* will also include an agent inhibitory to the growth of the *Agrobacterium* cells. Commonly used compounds are antibiotics such as cefotaxime and carbenicillin. The media used in the selection may be formulated to maintain transformed callus or suspension culture cells in an undifferentiated state, or to allow production of shoots from callus, leaf or stem segments, tuber disks, and the like.

Cells or callus observed to be growing in the presence of normally inhibitory concentrations of the selective agents are presumed to be transformed and may be subcultured several additional times on the same medium to remove nonresistant sections. The cells or callus can then be assayed for the presence of the viral gene cassette, or can be subjected to known plant regeneration protocols. In protocols involving the direct production of shoots, those shoots appearing on the selective media are presumed to be transformed and can be excised and rooted, either on selective medium suitable for the production of roots, or by simply dipping the excised shoot in a root-inducing compound and directly planting it in vermiculite.

In order to produce transgenic plants exhibiting viral resistance, the viral genes must be taken up into the plant cell and stably integrated within the plant genome. Plant cells and tissues selected for their resistance to an inhibitory agent are presumed to have acquired the selectable marker gene encoding this resistance during the transformation treatment. Since the marker gene is commonly linked to the viral genes,

it can be assumed that the viral genes have similarly been acquired. Southern blot hybridization analysis using a probe specific to the viral genes can then be used to confirm that the foreign genes have been taken up and integrated into the genome of the plant cell. This technique may also give some indication of the number of copies of the gene that have been incorporated. Successful transcription of the foreign gene into mRNA can likewise be assayed using Northern blot hybridization analysis of total cellular RNA and/or cellular RNA that has been enriched in a polyadenylated region. mRNA molecules encompassed within the scope of the invention are those which contain viral specific sequences derived from the viral genes present in the transformed vector which are of the same polarity as that of the viral genomic RNA such that they are capable of base pairing with viral specific RNA of the opposite polarity to that of viral genomic RNA under conditions described in Chapter 7 of Sambrook et al. (1989). Moreover, mRNA molecules encompassed within the scope of the invention are those which contain viral specific sequences derived from the viral genes present in the transformed vector which are of the opposite polarity to that of the viral genomic RNA such that they are capable of base pairing with viral genomic RNA under conditions described in Chapter 7 in Sambrook et al. (1989).

The presence of a viral coat protein gene can also be detected by immunological assays, such as the double-antibody sandwich assays described by Namba et al., Gene, 107, 181 (1991) as modified by Clark et al., J. Gen. Virol., 34, 475 (1979). See also, Namba et al., Phytopathology, 82, 940 (1992). Potyvirus resistance can also be assayed via infectivity studies as generally disclosed by Namba et al., ibid., wherein

plants are scored as symptomatic when any inoculated leaf shows veinclearing, mosaic or necrotic symptoms.

Seed from plants regenerated from tissue culture is grown in the field and self-pollinated to generate true breeding plants. The progeny from these plants become true breeding lines which are evaluated for viral resistance in the field under a range of environmental conditions. The commercial value of viral-resistant plants is greatest if many different hybrid combinations with resistance are available for sale. Additionally, hybrids adapted to one part of a country are not adapted to another part because of differences in such traits as maturity, disease and insect tolerance. Because of this, it is necessary to breed viral resistance into a large number of parental lines so that many hybrid combinations can be produced.

Adding viral resistance to agronomically elite lines is most efficiently accomplished when the genetic control of viral resistance is understood. This requires crossing resistant and sensitive plants and studying the pattern of inheritance in segregating generations to ascertain whether the trait is expressed as dominant or recessive, the number of genes involved, and any possible interaction between genes if more than one are required for expression. With respect to transgenic plants of the type disclosed herein, the transgenes exhibit dominant, single gene Mendelian behavior. This genetic analysis can be part of the initial efforts to convert agronomically elite, yet sensitive lines to resistant lines. A conversion process (backcrossing) is carried out by crossing the original transgenic resistant line with a sensitive elite line and crossing the progeny back to the sensitive parent. The progeny from this cross will segregate such that some plants carry the resistance gene(s) whereas some do not.

Plants carrying the resistance gene(s) will be crossed again to the sensitive parent resulting in progeny which segregate for resistance and sensitivity once more. This is repeated until the original sensitive parent has been converted to a resistant line, yet possesses all of the other important attributes originally found in the sensitive parent. A separate backcrossing program is implemented for every sensitive elite line that is to be converted to a virus resistant line.

Subsequent to the backcrossing, the new resistant lines and the appropriate combinations of lines which make good commercial hybrids are evaluated for viral resistance, as well as for a battery of important agronomic traits. Resistant lines and hybrids are produced which are true to type of the original sensitive lines and hybrids. This requires evaluation under a range of environmental conditions under which the lines or hybrids will be grown commercially. Parental lines of hybrids that perform satisfactorily are increased and utilized for hybrid production using standard hybrid production practices.

The invention will be further described by reference to the following detailed examples. Enzymes were obtained from commercial sources and were used according to the vendor's recommendations or other variations known in the art. Other reagents, buffers, etc., were obtained from commercial sources, such as GIBCO-BRL, Bethesda, MD, and Sigma Chemical Co., St. Louis, MO, unless otherwise specified.

Most of the recombinant DNA methods employed in practicing the present invention are standard procedures, well known to those skilled in the art, and described in detail in, for example, in European Patent

Application Publication Number 223,452, published November 29, 1986, which is incorporated herein by reference. General references containing such standard techniques include the following: R. Wu, ed., Methods in Enzymology, Vol. 68 (1979); J.H. Miller, Experiments in Molecular Genetics (1972); J. Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd ed. (1989); and D.M. Glover, ed., DNA Cloning Vol. II (1982). Figures 6 and 7 are presented to illustrate constructions of this invention.

Example I

A. Isolation of FLA83 W RNAs

Crookneck squash plants (7-days old) were inoculated with PRV strain W (watermelon) Florida-83; after 21 days, infected leaves were harvested and PRV virus particles were isolated. The procedure used was based on protocols from Purcifull et al. (Phytopathology, 69, 112 (1979)) for PRV type W isolation. Approximately 50 g of fresh leaf tissue were extracted in 100 mL of 0.5 M potassium phosphate buffer (pH 7.5, "PB") containing 0.1% sodium sulfate, 25 mL of chloroform, and 25 mL of carbon tetrachloride. After centrifugation of the extract at 1000 x g for 5 minutes, the pellet was resuspended in 50 mL of PB buffer and centrifuged again at 1000 x g for 5 minutes. The supernatants from each centrifugation were pooled, then centrifuged at 13,000 x g for 15 minutes. To the supernatant was added TRITON X-100 to a final concentration of 1% (v/v), polyethylene glycol (PEG) 8,000 (Research Grade from Sigma Chemical Co.) to a final concentration of 4%, (w/v) and NaCl to a final concentration of 100 mM. The suspension was stirred for 1 hour at 0-4°C. It was then centrifuged at 10,000 x g for 10 minutes.

The virus pellet was collected and resuspended in about 40 mL of PB buffer. After centrifugation at 12,000 x g for 10 minutes, the pellet was discarded and virus was precipitated by adding PEG to a final concentration of 8% (w/v) and NaCl to a final concentration of 100 mM and stirring for 0.5 hour at 0-4°C. After centrifugation at 12,000 x g for 10 minutes the pellets were resuspended with the aid of a tissue grinder in 5 mL of 20 mM PB buffer and layered over a 30% Cs_2SO_4 cushion. This was centrifuged in a Beckman Ti75 at 140,000 x g for 18 hours at 5°C. After centrifugation, the virus band was harvested, and dialyzed against 20 mM PB buffer overnight at 4°C. The dialyzed virus preparation was lysed and viral RNA precipitated with LiCl (2 M final concentration). The viral RNA was recovered by centrifugation. Viral RNA was dissolved and precipitated by ethanol and resuspended in water.

B. Cloning and Engineering PRV FLA83 Coat Protein Gene

To obtain engineered genes of the PRV strain FLA83 coat protein gene, the following steps were carried out: 1) single-stranded cDNA of PRV FLA83 was constructed; 2) coat protein sequences were amplified by PCR; 3) the PRV CP PCR product was cloned; 4) expression cassettes were inserted into binary vectors; 5) plants transgenic for the PRV CP construct were produced; and 6) progeny of R_0 transgenic plants were challenged to identify protected lines.

Single-stranded cDNA of PRV FLA83 W RNA was synthesized with the use of ClonStruct™ cDNA Library Construction Kit reagents (US Biochemical, Cleveland, OH). Briefly, a first strand cDNA synthesis reaction was primed with the vector primer pTRXN PLUS (US Biochemical, Cleveland, OH). This vector includes a poly dT tract; the plasmid poly dT tract anneals with the poly A^+ tail

of PRV RNA. Subsequently, the PRV first strand cDNA was synthesized; the reaction extended the pTRXN plasmid primer.

PRV single-stranded cDNA was used as a template to PCR amplify PRV coat protein sequences. Two versions of the coat protein coding sequence were amplified: long (primers used were RMM384 and RMM385) and a short version (primers used were RMM388 and RMM385) (Figure 3 [SEQ ID NOS:5, 6 and 7 for RMM384, RMM385, and RMM 388, respectively]). Sequences for NcoI sites were included in each of these primers, so that the PCR products contained NcoI sites which were generated during the amplification. After amplification, coat protein gene PCR products were digested with NcoI in preparation for insertion into the NcoI site of pUC18cpexpress. Both the long and short versions were installed into pUC18cpexpress. The long PRV FLA83 CP gene in cpexpress is known as FLA83CPpUC18cpexp16 (Figure 6); the short PRV FLA83 CP gene in cpexpress is known as FLA83CPpUC18cpexp34 (Figure 7).

The CP coding sequences of each were then nucleotide sequenced with the use of USB Sequenase Version II sequencing Kit (Figures 1 and 2 [SEQ ID NOS:1 and 3]). The coat protein gene sequence of the FLA83 PRV strain is novel information. Comparison with the coat protein genes of 5 different PRV strains shows that the CP gene of FLA83 differs from characterized coat protein sequences of other PRV strains in at least 15 amino acid positions (Figure 5).

After insertion into the expression cassette pUC18cpexpress, both sense and antisense cassettes were obtained (Figures 6 and 7). Subsequently, HindIII fragments harboring FLA83CPpUC18cpexp16 sense or antisense and FLA83CPpUC18cpexp34 sense or antisense

were isolated and installed into the plasmid pUC1318 (Kay et al., Nuc. Acids Res., 15:2778 (1987)) to provide additional cloning sites for insertion into binary vectors. Both sense and antisense versions of the long and short PRV FLA83 cassettes were excised as BamHI fragments and installed into the BglII site of binary plasmids. FLA83 coat protein expression cassettes were inserted in combination with other coat protein cassettes in binary vectors as summarized below in Table 1:

Table 1

Binary	Parental Plasmid	Site	FLA83 CP Used	pEPG#
pGA482G	pEPG192 (V27cp)	XbaI	Short pUC1318cpexp34(s)	194
pGA482G	pEPG191 (V27cp)	XbaI	Long pUC1318cpexp16(s)	241
pGA482G	pEPG198 (V33cp)	XbaI	Short pUC1318cpexp34(s)	242
pGA482G	pEPG198 (V33cp)	XbaI	Long pUC1318cpexp16(s)	249
pPRBN	pEPG111 (CZW)	BglII	Long pUC18cpexp16(s)	208 or 252
pPRBN	pEPG111 (CZW)	BglII	Long pUC18cpexp16(as)	207
pPRBN	pEPG111 (CZW)	BglII	Short pUC18cpexp34(s)	209
pPRBN	pEPG111 (CZW)	BglII	Short pUC18cpexp34(as)	210
pPRBN	pEPG109 (CwlZW)	BglII	Long pUC18cpexp16(s)	212 or 253
pPRBN	pEPG109 (CwlZW)	BglII	Long pUC18cpexp16(as)	211
pPRBN	pEPG109 (CwlZW)	BglII	Short pUC18cpexp34(s)	213
pPRBN	pEPG109 (CwlZW)	BglII	Short pUC18cpexp34(as)	214
pGA482G	pEPG189 (CMV-C)	BglII	Long pUC18cpexp16(s)	216
pGA482G	pEPG189 (CMV-C)	BglII	Long pUC18cpexp16(as)	215
pGA482G	pEPG189 (CMV-C)	BglII	Short pUC18cpexp34(s)	218
pGA482G	pEPG189 (CMV-C)	BglII	Short pUC18cpexp34(as)	220
pGA482G	pEPG120 (Cwl62)	BglII	Short pUC18cpexp34(s)	222
pGA482G	pEPG120 (Cwl62)	BglII	Short pUC18cpexp34(as)	223
pGA482G	pEPG120 (Cwl62)	BglII	Long pUC18cpexp16(s)	236

pPRBN	pEPG106 (ZW)	HindIII	Long pUC18cpexp16(as)	203
pPRBN	pEPG106 (ZW)	HindIII	Long pUC18cpexp16(s)	204
pPRBN	pEPG106 (ZW)	HindIII	Short pUc18cpexp34(s)	205
pPRBN	pEPG106 (ZW)	HindIII	Short pUC18cpexp34(as)	206
pGA482G	pEPG321 (SqBV)	HpaI	Short pUC18cpexp34(s)	327*
pGA482G	pEPG321 (SqBV)	HpaI	Long pUC18cpexp16(s)	328#

#A BsrBI fragment, including all of the CP cassettes found in pEPG212, was isolated from pEPG212 and installed into the HpaI site of pEPG321 to give pEPG328.

*A BsrBI fragment, including all the CP cassettes found in pEPG213, was isolated from pEPG213 and installed into the HpaI site of pEPG321 to give pEPG327.

For further information on CMV-C and CMV-wl see Quemada et al., J. Gen. Virol., 70, 1065 (1989). For further information on CMV V27 and V33 coat proteins, see Applicants' Assignees copending Patent Application Serial No. 08/367,789 entitled "Plants Resistant to V27, V33, or V34 Strains of Cucumber Mosaic Virus" filed on December 30, 1994, incorporated by reference herein. For further information on ZYMV and WMVII coat protein genes see Applicants' Assignees copending Patent Application Serial No. 08/232,846 entitled "Potyvirus Coat Protein Genes and Plants Transformed Therewith" filed on April 25, 1994, incorporated by reference herein. For further information on SqBV coat proteins see Applicants' Assignees copending Patent Application Serial No. 08/085,250 entitled "Squash Mosaic Virus Genes and Plants Transformed Therewith" filed on June 30, 1993, incorporated by reference herein.

Agrobacterium-mediated transfer of the plant expressible PRV coat protein genes described herein was done using the methods described in PCT published application WO 89/05859, entitled "*Agrobacterium* Mediated Transformation of Germinating Plant Seeds." Binary plasmids listed above (for further information, refer to Applicants' Assignees copending Patent Application Serial No. 08/366,991 entitled "Transgenic Plants Expressing DNA Constructs Containing a Plurality of Genes to Impart Virus Resistance" filed on December 30, 1994, and incorporated by reference herein) were transformed into the *A. tumefaciens* strains C58Z707 (obtained from Dr. A.G. Hepburn, University of Illinois, Urbana, Illinois) and Mog301 (obtained from Mogen NV, Leiden, The Netherlands). The resulting *Agrobacteria* strains have been used for plant transformations.

C. Cloning, Sequencing, and Engineering PRV Brazil Coat Protein Gene

A virion preparation of PRV Brazil isolate was prepared by Dr. Gonsalves. Subsequently, virion RNA and reverse transcribed signal stranded cDNA were isolated. Coat protein sequences were amplified by PCR; 5' and 3' terminal NcoI sites were installed during the PCR to amplify the coat protein sequence as described by the "proteolytic sites" described in Quemada et al. (1990). Subsequently, the NcoI fragment obtain by PCR amplification was cloned into pGMM (derived from phagemid pBLUESCRIPT II SK (+) [Strategene, La Jolla, CA] to obtain pGMM/PRV-Brcp-7-2.

The CP coding sequence of pGMM/PRV-Brcp-7-2 was nucleotide sequenced with the use of US Biochemical (Cleveland, OH) Sequenase Version II sequencing Kit (Figure 8). The predicted coat protein sequence of the PRV Brazil isolate is shown in Fig. 9. Comparison with

the coat protein genes of seven different PRV strains shows that the CP gene of Brazil PRV differed from characterized coat protein sequences of other PRV strains (Figures 8, 9, and 10).

After the CP gene was sequenced, NcoI fragments were prepared and inserted into the NcoI site of the expression cassette pUC1318cpexpress. Subsequently, HindIII fragments harboring PRV Brazil coat protein sense cassette were inserted into the HindIII site of the binary plasmid pGA482G. The resulting binary plasmid was transformed into *A. tumefaciens* strains C58Z707 and Mog301. The resulting *Agrobacteria* strains were used for plant transformations.

All publications, patents and patent documents are incorporated by reference herein, as though individually incorporated by reference. The invention has been described with reference to various specific and preferred embodiments and techniques. However, it should be understood that many variations and modifications may be made while remaining within the spirit and scope of the invention.

WHAT IS CLAIMED IS:

1. An isolated and purified DNA molecule consisting essentially of DNA encoding the coat protein of the FLA83 W strain of papaya ringspot virus.

2. The isolated and purified DNA molecule of claim 1 from the FLA83 W strain of papaya ringspot virus having the nucleotide sequence shown in Figure 1 [SEQ ID NO:1].

3. A vector comprising a chimeric expression cassette comprising the DNA molecule of claim 1, a promoter and a polyadenylation signal, wherein the promoter is operably linked to the DNA molecule, and the DNA molecule is operably linked to the polyadenylation signal.

4. The vector of claim 3 wherein the promoter is the cauliflower mosaic virus 35S promoter.

5. The vector of claim 4 wherein the polyadenylation signal is the polyadenylation signal of the cauliflower mosaic 35S gene.

6. A bacterial cell comprising the vector of claim 3.

7. The bacterial cell of claim 6 wherein the bacterial cell is selected from the group consisting of an *Agrobacterium tumefaciens* cell and an *Agrobacterium rhizogenes* cell.

8. A transformed plant cell transformed with the vector of claim 3.

9. The transformed plant cell of claim 8 wherein the promoter is cauliflower mosaic virus 35S promoter and

the polyadenylation signal is the polyadenylation signal of the cauliflower mosaic 35S gene.

10. A plant selected from the family *Cucurbitaceae* comprising a plurality of the transformed cells of claim 8.

11. The isolated and purified DNA molecule of claim 1 from the FLA83 strain of papaya ringspot virus having the nucleotide sequence shown in Figure 2 [SEQ ID NO:3].

12. A method of preparing a papaya ringspot viral resistant plant comprising:

(a) transforming plant cells with a chimeric expression cassette comprising a promoter functional in plant cells operably linked to a DNA molecule that encodes a coat protein; wherein the DNA molecule is derived from a papaya ringspot virus strain FLA83 W;

(b) regenerating the plant cells to provide a differentiated plant; and

(c) identifying a transformed plant that expresses the papaya ringspot virus coat protein at a level sufficient to render the plant resistant to infection by the papaya ringspot virus strain.

13. The method of claim 12 wherein the DNA molecule is derived from a papaya ringspot virus strain having the nucleotide sequence shown in Figure 1 [SEQ ID NO:1] or Figure 2 [SEQ ID NO:3].

14. The method of claim 12 wherein the dicot is selected from the family *Cucurbitaceae*.

15. A vector comprising a chimeric expression cassette comprising the DNA molecule of claim 1 and at least one chimeric expression cassette comprising a heterologous

PRVcoat protein gene, a cucumber mosaic virus coat protein gene, a squash mosaic virus coat protein gene, a zucchini yellow mosaic virus coat protein gene, or a watermelon mosaic virus-2 coat protein gene, wherein each expression cassette comprises a promoter and a polyadenylation signal wherein the promoter is operably linked to the DNA molecule, and the DNA molecule is operably linked to the polyadenylation signal.

16. A bacterial cell comprising the vector of claim 15.

17. A transformed plant cell transformed with the vector of claim 15.

18. The transformed plant cell of claim 17 wherein the promoter is cauliflower mosaic virus 35S promoter and the polyadenylation signal is the polyadenylation signal of the cauliflower mosaic 35S gene.

FIG. 1A

```
1   ATGGCTCCATTCAATGAGCTGGCGAAACAAGGGAGGGCCCCATACGTCTCGGAAGTTGGATTAAGAAGGTTGTATACGTGTGAACGCGGATCAGTGGATG   100
    MetAlaProPheAsnGluLeuAlaLysGlnGlyArgAlaProTyrValSerGluValGlyLeuArgArgLeuTyrThrCysGluArgGlySerValAspG
    M  A  P  F  N  E  L  A  K  Q  G  R  A  P  Y  V  S  E  V  G  L  R  R  L  Y  T  C  E  R  G  S  V  D  E

101 AATTGGAAGCGTATATAGATAAATATTTTGAGCGTGAGAGGGGAGACTCACCCGAAGTACTGGTGTACCATGAATCAAGGAGTACTGATGATTATGAACT   200
    luLeuGluAlaTyrIleAspLysTyrPheGluArgGluArgGlyAspSerProGluValLeuValTyrHisGluSerArgSerThrAspAspTyrGluLe
      L  E  A  Y  I  D  K  Y  F  E  R  E  R  G  D  S  P  E  V  L  V  Y  H  E  S  R  S  T  D  D  Y  E  L

201 TGTTCGTGTCAACAATACACATGTGTTTCATCAAGCCAAAAATGAAGCTGTGGACGCTGGTTTGAACGAAAAGCTCAAAGAAAAAGAAAAACAGAGAGAG   300
    uValArgValAsnAsnThrHisValPheHisGlnAlaLysAsnGluAlaValAspAlaGlyLeuAsnGluLysLeuLysGluLysGluLysGlnArgGlu
      V  R  V  N  N  T  H  V  F  H  Q  A  K  N  E  A  V  D  A  G  L  N  E  K  L  K  E  K  E  K  Q  R  E

301 AAAGAAAAAGAAAAACAAAAAGAGAAAGAAAAAGATGATGCTAGTGACGGAAATGATGTGTTAACTAGCACAAAAACTGGAGAGAGAGATAGAGATGTCA   400
    LysGluLysGluLysGlnLysGluLysGluLysAspAspAlaSerAspGlyAsnAspValLeuThrSerThrLysThrGlyGluArgAspArgAspValA
    K  E  K  E  K  Q  K  E  K  E  K  D  D  A  S  D  G  N  D  V  L  T  S  T  K  T  G  E  R  D  R  D  V  N

401 ATGTTGGAACTAGTGGGACTTTCACTATTCCAAGGATTAAACCATTCAATGATAAGATGATTTTACCGAGAATTAAGGGAAAAACTGTCCTTAATTTAAA   500
    snValGlyThrSerGlyThrPheThrIleProArgIleLysProPheAsnAspLysMetIleLeuProArgIleLysGlyLysThrValLeuAsnLeuAs
      V  G  T  S  G  T  F  T  I  P  R  I  K  P  F  N  D  K  M  I  L  P  R  I  K  G  K  T  V  L  N  L  N

501 TCATCTTCTTCAGTATAATCCGCAACAAATTGACATTTCGAACACTCGTGCCACTCAGTCACAATTTGAAAAATGGCACGAGGGAGTGAGGAATGATTAT   600
    nHisLeuLeuGlnTyrAsnProGlnGlnIleAspIleSerAsnThrArgAlaThrGlnSerGlnPheGluLysTrpHisGluGlyValArgAsnAspTyr
      H  L  L  Q  Y  N  P  Q  Q  I  D  I  S  N  T  R  A  T  Q  S  Q  F  E  K  W  H  E  G  V  R  N  D  Y

601 GGCCTGAATGATAAAGAGATGGAAGTAATGTTAAATGGCTTGATGGTTTGGTGTATTGAGAATGGTACATCTCCGGACATATCTGGTGTCTGGGTTATGA   700
    GlyLeuAsnAspLysGluMetGluValMetLeuAsnGlyLeuMetValTrpCysIleGluAsnGlyThrSerProAspIleSerGlyValTrpValMetM
      G  L  N  D  K  E  M  E  V  M  L  N  G  L  M  V  W  C  I  E  N  G  T  S  P  D  I  S  G  V  W  V  M  M
```

FIG. 1B

```
701  TGGATGATACTACAGGAACCCAAGTTGATTATCCAATCAAGCCTTTAATTGAGCATGCTACTCCGTCATTTAGGCAAATTATGGCTCACTTTAGTAACGC  800
     etAspAspThrThrGlyThrGlnValAspTyrProIleLysProLeuIleGluHisAlaThrProSerPheArgGlnIleMetAlaHisPheSerAsnAl
       D  D  T  T  G  T  Q  V  D  Y  P  I  K  P  L  I  E  H  A  T  P  S  F  R  Q  I  M  A  H  F  S  N  A

801  GGCAGAAGCATACATTGCAAGGAGAAATGCTACTGAGAGGTACATGCCGCGGTATGGAATCAAGAGAAATTTGACTGACATTAGCCTCGCTAGATACGCT  900
     aAlaGluAlaTyrIleAlaArgArgAsnAlaThrGluArgTyrMetProArgTyrGlyIleLysArgAsnLeuThrAspIleSerLeuAlaArgTyrAla
       A  E  A  Y  I  A  R  R  N  A  T  E  R  Y  M  P  R  Y  G  I  K  R  N  L  T  D  I  S  L  A  R  Y  A

901  TTCGATTTCTATGAGGTTAATTCGAAAACACCTGATAGGGCTCGCGAAGCTCGCATGCAGATGAAAGCTGCAGCGCTGCGAAACACTAATCGCAGAATGT  1000
     PheAspPheTyrGluValAsnSerLysThrProAspArgAlaArgGluAlaArgMetGlnMetLysAlaAlaAlaLeuArgAsnThrAsnArgArgMetP
       F  D  F  Y  E  V  N  S  K  T  P  D  R  A  R  E  A  R  M  Q  M  K  A  A  A  L  R  N  T  N  R  R  M  F

1001 TTGGTATGGACGGCAGTGTCAGTAACAAGGAAGAAAATACGGAGAGACACACAGTGGAAGATGTCAATAGAGACATGCACTCTCTCCTGGGTATGCGCAA  1100
     heGlyMetAspGlySerValSerAsnLysGluGluAsnThrGluArgHisThrValGluAspValAsnArgAspMetHisSerLeuLeuGlyMetArgAs
       G  M  D  G  S  V  S  N  K  E  E  N  T  E  R  H  T  V  E  D  V  N  R  D  M  H  S  L  L  G  M  R  N

1101 CTGAATACTCGCGCTTGTGTGTTTGTCGAGTCTAACTCGACCCTGTTTCACCCCATGG  1158
     nEndIleLeuAlaLeuValCysLeuSerSerLeuThrArgProCysPheThrProTrp
       *  I  L  A  L  V  C  L  S  S  L  T  R  P  C  F  T  P  W
```

FIG. 2A

```
1   CCATGGCCAAGAATGAAGCTGTGGACGCTGGTTTGAACGAAAAGCTCAAAGAAAAAGAAAAACAGAGAGAGAAAGAAAAAGAAAAACAAAAAGAGAAAGA   100
      MetAlaLysAsnGluAlaValAspAlaGlyLeuAsnGluLysLeuLysGluLysGluLysGlnArgGluLysGluLysGluLysGlnLysGluLysGl
      M  A  K  N  E  A  V  D  A  G  L  N  E  K  L  K  E  K  E  K  Q  R  E  K  E  K  E  K  Q  K  E  K  E

101 AAAAGATGATGCTAGTGACGGAAATGATGTGTTAACTAGCACAAAAACTGGAGAGAGAGATAGAGATGTCAATGTTGGAACTAGTGGGACTTTCACTATT   200
      uLysAspAspAlaSerAspGlyAsnAspValLeuThrSerThrLysThrGlyGluArgAspArgAspValAsnValGlyThrSerGlyThrPheThrIle
      K  D  D  A  S  D  G  N  D  V  L  T  S  T  K  T  G  E  R  D  R  D  V  N  V  G  T  S  G  T  F  T  I

201 CCAAGGATTAAACCATTCAATGATAAGATGATTTTACCGAGAATTAAGGGAAAAACTGTCCTTAATTTAAATCATCTTCTTCAGTATAATCCGCAACAAA   300
      ProArgIleLysProPheAsnAspLysMetIleLeuProArgIleLysGlyLysThrValLeuAsnLeuAsnHisLeuLeuGlnTyrAsnProGlnGlnI
      P  R  I  K  P  F  N  D  K  M  I  L  P  R  I  K  G  K  T  V  L  N  L  N  H  L  L  Q  Y  N  P  Q  Q  I

301 TTGACATTTCGAACACTCGTGCCACTCAGTCACAATTTGAAAAATGGCACGAGGGAGTGAGGAATGATTATGGCCTGAATGATAAAGAGATGGAAGTAAT   400
      leAspIleSerAsnThrArgAlaThrGlnSerGlnPheGluLysTrpHisGluGlyValArgAsnAspTyrGlyLeuAsnAspLysGluMetGluValMe
        D  I  S  N  T  R  A  T  Q  S  Q  F  E  K  W  H  E  G  V  R  N  D  Y  G  L  N  D  K  E  M  E  V  M

401 GTTAAATGGCTTGATGGTTTGGTGTATTGAGAATGGTACATCTCCGGACATATCTGGTGTCTGGGTTATGATGGATGATACTACAGGAACCCAAGTTGAT   500
      tLeuAsnGlyLeuMetValTrpCysIleGluAsnGlyThrSerProAspIleSerGlyValTrpValMetMetAspAspThrThrGlyThrGlnValAsp
        L  N  G  L  M  V  W  C  I  E  N  G  T  S  P  D  I  S  G  V  W  V  M  M  D  D  T  T  G  T  Q  V  D

501 TATCCAATCAAGCCTTTAATTGAGCATGCTACTCCGTCATTTAGGCAAATTATGGCTCACTTTAGTAACGCGGCAGAAGCATACATTGCAAGGAGAAATG   600
      TyrProIleLysProLeuIleGluHisAlaThrProSerPheArgGlnIleMetAlaHisPheSerAsnAlaAlaGluAlaTyrIleAlaArgArgAsnA
      Y  P  I  K  P  L  I  E  H  A  T  P  S  F  R  Q  I  M  A  H  F  S  N  A  A  E  A  Y  I  A  R  R  N  A
```

FIG. 2B

```
601  CTACTGAGAGGTACATGCCGCGGTATGGAATCAAGAGAAATTTGACTGACATTAGCCTCGCTAGATACGCTTTCGATTTCTATGAGGTTAATTCGAAAAC
     laThrGluArgTyrMetProArgTyrGlyIleLysArgAsnLeuThrAspIleSerLeuAlaArgTyrAlaPheAspPheTyrGluValAsnSerLysTh
       T  E  R  Y  M  P  R  Y  G  G  I  K  R  N  L  T  D  I  S  L  A  R  Y  A  F  D  F  Y  E  V  N  S  K  T

701  ACCTGATAGGGCTCGCCGAAGCTCACATGCAGATGAAAGCTGCAGCGCTGCGAAACACTAATCGCAGAATGTTTGGTATGGACGGCAGTGTCAGTAACAAG
     rProAspArgAlaArgGluAlaHisMetGlnMetLysAlaAlaAlaLeuArgAsnThrAsnArgArgMetPheGlyMetAspGlySerValSerAsnLys
       P  D  R  A  R  E  A  H  M  Q  M  K  A  A  A  A  L  R  N  T  N  R  R  M  F  G  M  D  G  S  V  S  N  K

801  GAAGAAAATACGGAGAGACACACAGTGGAAGATGTCAATAGACACATGCCACTCTCTCCTGGGTATGCGCAACTGAATACTCGCGCTTGTGTGTTTGTCGA
     GluGluAsnThrGluArgHisThrValGluAspValAsnArgAspMetHisSerLeuLeuGlyMetArgAsnEndIleLeuAlaLeuValCysLeuSerS
       E  E  N  T  E  R  H  T  V  E  D  V  N  R  D  M  H  S  L  L  G  M  R  N  *  I  L  A  L  V  C  L  S  S

901  GTCTAACTCGACCCTGTTTCACCCCATGG  929
     erLeuThrArgProCysPheThrProTrp
       L  T  R  P  C  F  T  P  W
```

FIG. 3A

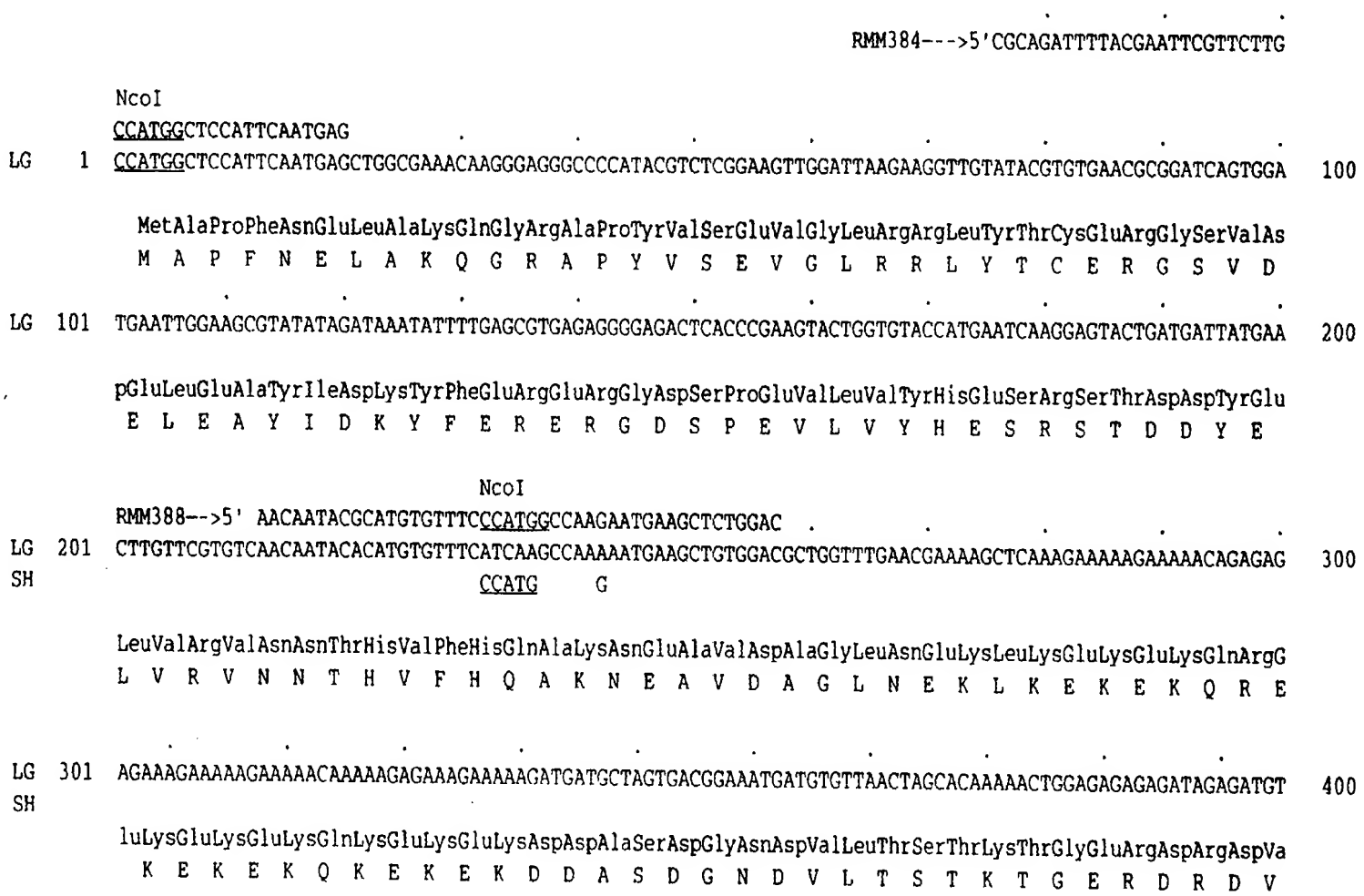


FIG. 3B

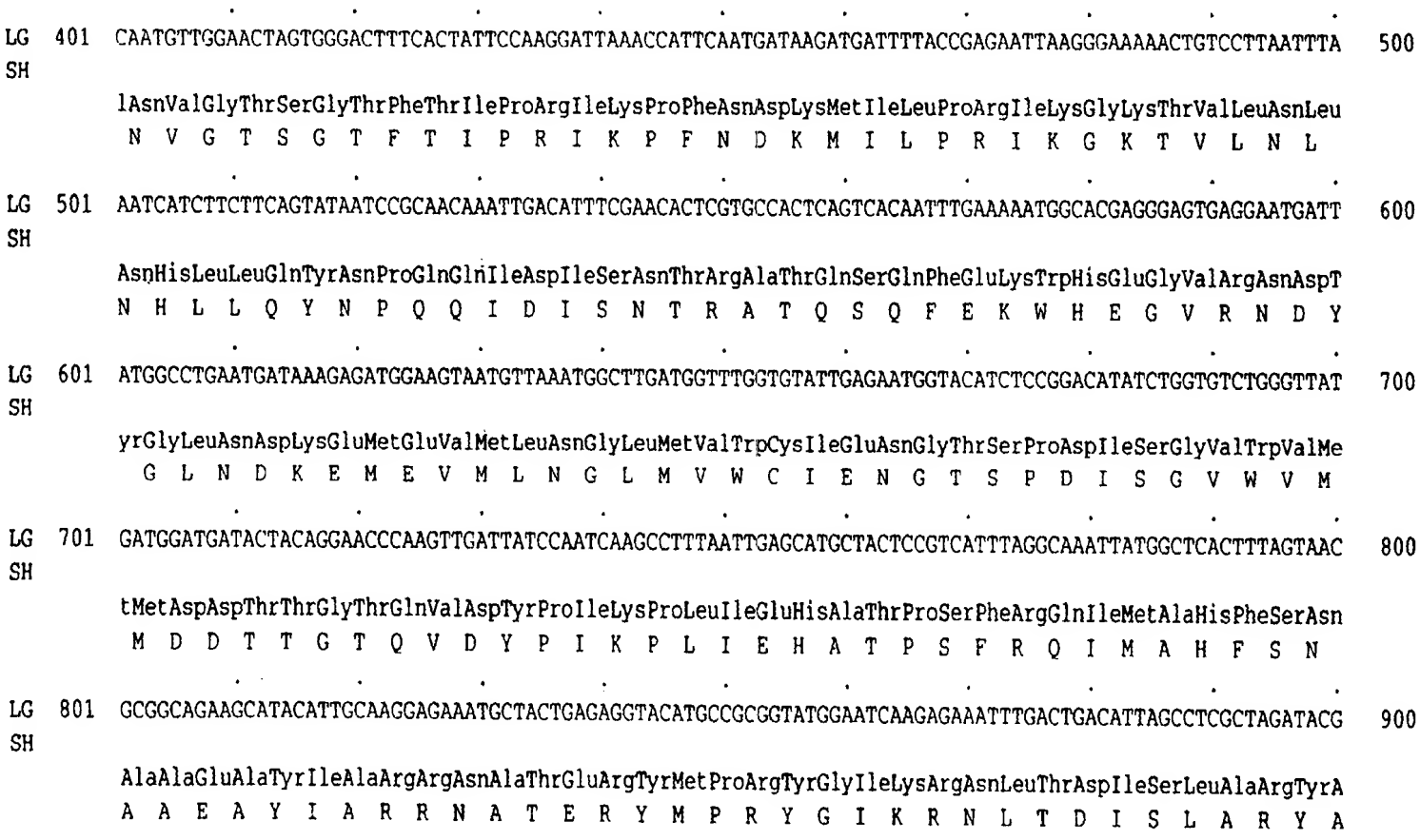


FIG. 3C

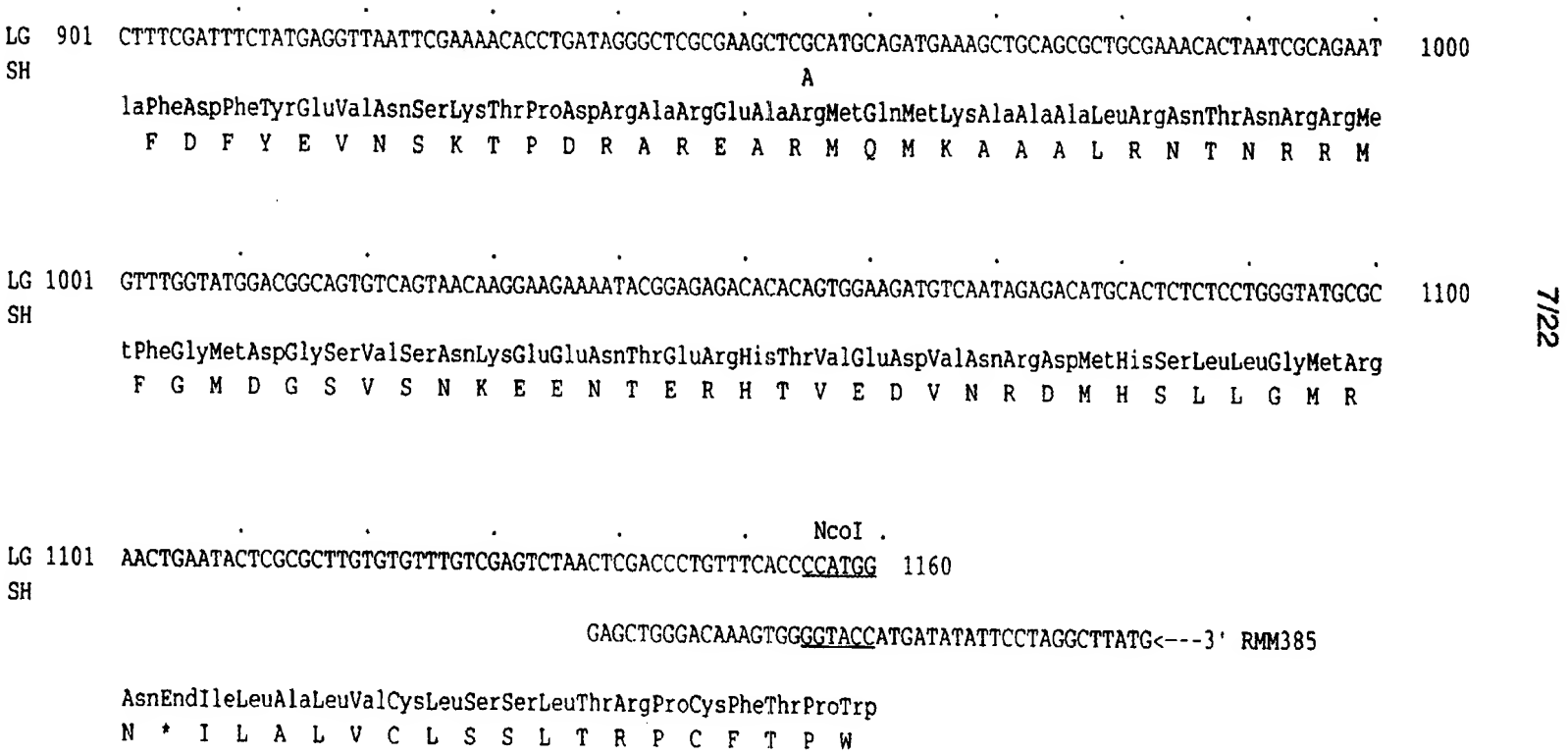


FIG. 4A

```
               201                                                                                                              300
Australiancp-W .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
        Hacp-P GATTTTACCA ATGGGTTCTT GAGCAAGCTC CATTCAATGA GTTGGCGAAA CAAGGAAGGG CCCCATACGT CTCGGAAGTT GGATTAAGAA GATTGTACAC
       Usacp-P GATTTTACCA ATGGGTTCTT GAGCAAGCTC CATTCAATGA GTTGGCGAAA CAAGGAAGGG CCCCATACGT CTCGGAAGTT GGATTAAGAA GATTGTACAC
       Usacp-W GATTTTACCA ATGGGTTCTT GAGCAAGCTC CATTCAATGA GTTGGCGAAA CAAGGAAGGG CCCCATACGT CTCGGAAGTT GGATTAAGAA GATTGTACAC
     Fla831cp-W .......... .......... .CCATGGCTC CATTCAATGA GCTGGCGAAA CAAGGGAGGG CCCCATACGT CTCGGAAGTT GGATTAAGAA GGTTGTATAC

               301                                                                                                              400
Australiancp-W .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
        Hacp-P AAGTGAACGT GGATCAATGG ACGAATTAGA AGCGTATATA GATAAATACT TTGAGCGTGA GAGAGGAGAC TCGCCCGAAT TACTAGTGTA CCATGAATCA
       Usacp-P AAGTGAACGT GGATCAATGG ACGAATTAGA AGCGTATATA GATAAATACT TTGAGCGTGA GAGAGGAGAC TCGCCCGAAT TACTAGTGTA CCATGAATCA
       Usacp-W AAGTGAACGT GGATCAATGG ATGAATTAGA AGCGTATATA GATAAATACT TTGAGCGTGA GAGAGGAGAC TCACCCGAAT TACTAGTGTA CCATGAATCA
     Fla831cp-W GTGTGAACGC GGATCAGTGG ATGAATTGGA AGCGTATATA GATAAATATT TTGAGCGTGA GAGGGGAGAC TCACCCGAAG TACTGGTGTA CCATGAATCA

               401                                                                                                              500
Australiancp-W .......... .......... .......... .......... .......... .......TCC AAAAATGAAG CTGTGGATGC TGGTTTGAAC GAAAAGCTCA
        Hacp-P AGGGGCACTG ATGATTATCA ACTTGTTTGT AGCAACAATA CGCATGTGTT TCATCAGTCC AAGAATGAAG CTGTGGATGC TGGTTTGAAT GAAAAACTCA
       Usacp-P AGGAGCACTG ATGATTATCA ACTTGTTTGT AGCAACAATA CGCATGTGTT TCATCAGTCC AAGAATGAAG CTGTGGATGC TGGTTTGAAT GAAAAACTCA
       Usacp-W AGGAGCACTG ATGATTATCA ACTTGTTTGC AGTAACAATA CACATGTGTT TCATCAGTCC AAAAATGAAG CTGTGGATAC TGGTTTGAAT GAAAAATTCA
     Fla831cp-W AGGAGTACTG ATGATTATGA ACTTGTTCGT GTCAACAATA CACATGTGTT TCATCAAGCC AAAAATGAAG CTGTGGACGC TGGTTTGAAC GAAAAGCTCA

               501                                                                                                              600
Australiancp-W AAGAAAAAGA AAAACAGAAA GAAAAAGAAA AAGAAAAACA AAAAGAGAAA GAGAAAGACG ATGCTAGTGA CGGAAATGAT GTGTCAACTA GCACAAAAAC
        Hacp-P AAGAGAAGGA AAAACAGAAA GAAAAAGAAA AAGAAAAACA AAAAGAGAAA GAAAAAGACG GTGCTAGTGA CGGAAATGAT GTGTCAACTA GCACAAAAAC
       Usacp-P AAGAGAAGGA AAATCAGAAA GAAAAAGAAA AAGAAAAACA AAAAGAGAAA GAAAAAGACG GTGCTAGTGA CGGAAATGAT GTGTCAACTA GCACAAAAAC
       Usacp-W AAGAAAAGGA AAAACAGAAA GAAAAAGAAA AAGAAAAACA AAAAGAGAAA GAAAAAGACG ATGCTAGTGA CGGAAATGAT GTGTCAACTA GCACAAAAAC
     Fla831cp-W AAGAAAAAGA AAAACAGAGA GAGAAAGAAA AAGAAAAACA AAAAGAGAAA GAAAAAGATG ATGCTAGTGA CGGAAATGAT GTGTTAACTA GCACAAAAAC
```

FIG. 4B

```
                601                                                                                                         700
Australiancp-W  TGGAGAGAGA GATAGAGATG TCAATGTTGG GACCAGTGGA ACTTTCACTG TTCCAAGAAT CAAATCATTT ACTGACAAGA TGATTCTACC AAGAATTAAG
        Hacp-P  TGGAGAGAGA GATAGAGATG TCAATGTTGG GACCAGTGGA ACTTTCACTG TTCCGAGAAT TAAATCATTT ACTGATAAGA TGGTTCTACC GAGAATTAAG
       Usacp-P  TGGAGAGAGA GATAGAGATG TCAATGTTGG GACCAGTGGA ACTTTCACTG TTCCGAGAAT TAAATCATTT ACTGATAAGA TGGTTCTACC GAGAATTAAG
       Usacp-W  TGGAGAGAGA GATAGAGATG TCAATGTTGG GACCAGTGGA ACTTTCACTG TTCCGAGAAT TAAATCATTT ACTGATAAGA TGATTCTACC GAGAATTAAG
     Fla831cp-W TGGAGAGAGA GATAGAGATG TCAATGTTGG AACTAGTGGG ACTTTCACTA TTCCAAGGAT TAAACCATTC AATGATAAGA TGATTTTACC GAGAATTAAG

                701                                                                                                         800
Australiancp-W  GGAAAGACTG TCCTTAATTT AAATCACCTT CTTCAGTATA ACCCGCAACA AATTGACATT TCTAACACTC GTGCCACTCA GTCACAATTT GAGAAGTGGT
        Hacp-P  GGGAAGACTG TCCTTAATTT AAATCATCTT CTTCAGTACA ATCCGCAACA AATTGACATT TCTAACACTC GTGCCACTCA TTCACAATTT GAGAAGTGGT
       Usacp-P  GGGAAGACTG TCCTTAATTT AAATCATCTT CTTCAGTACA ATCCGCAACA AATTGACATT TCTAACACTC GTGCCACTCA TTCACAATTT GAGAAGTGGT
       Usacp-W  GGAAAGTCTG TCCTTAATTT AAATCACCTA CTTCAGTATA ATCCGCAACA AATTGACATT TCTAACACTC GTGCCACTCA GTCACAATTT GAGAAGTGGT
     Fla831cp-W GGAAAAACTG TCCTTAATTT AAATCATCTT CTTCAGTATA ATCCGCAACA AATTGACATT TCGAACACTC GTGCCACTCA GTCACAATTT GAAAAATGGC

                801                                                                                                         900
Australiancp-W  ATGAGGGAGT GAGGAATGAT TATGGCCTTA ATGATAATGA AATGCAAGTG ATGCTAAATG GCTTGATGGT TTGGTGTATC GAGAATGGTA CATCTCCAGA
        Hacp-P  ATGAGGGAGT GAGGAATGAT TATGGCCTTA ATGATAATGA AATGCAAGTG ATGCTAAATG GTTTGATGGT TTGGTGTATC GAGAATGGTA CATCTCCAGA
       Usacp-P  ATGAGGGAGT GAGGAATGAT TATGGCCTTA ATGATAATGA AATGCAAGTG ATGCTAAATG GTTTGATGGT TTGGTGTATC GAGAATGGTA CATCTCCAGA
       Usacp-W  ATGAGGGAGT GAGGAATGAT TATGGCCTTA ATGATAATGA AATGCAAGTG ATGCTAAATG GTTTGATGGT TTGGTGTATC GAGAATGGTA CATCTCCAGA
     Fla831cp-W ACGAGGGAGT GAGGAATGAT TATGGCCTGA ATGATAAAGA GATGGAAGTA ATGTTAAATG GCTTGATGGT TTGGTGTATT GAGAATGGTA CATCTCCGGA

                901                                                                                                         1000
Australiancp-W  CATATCTGGT GTCTGGGTTA TGATGGATG. .....GGGAA ACCCAAGTTG ATTATCCAAT CAAGCCTTTA ATTGAGCATG CTACTCCGAC ATTTAGGCAA
        Hacp-P  CATATCTGGT GTCTGGGTTA TGATGGATG. .....GGGAA ACCCAAGTTG ATTATCCAAT CAAGCCTTTG ATTGAGCATG CTACTCCGTC ATTTAGGCAA
       Usacp-P  CATATCTGGT GTCTGGGTTA TGATGGATG. .....GGGAA ACCCAAGTTG ATTATCCAAT CAAGCCTTTG ATTGAGCATG CTACTCCGTC ATTTAGGCAA
       Usacp-W  CATATCTGGT GTCTGGGTTA TGATGGATG. .....GGGAA ACCCAAGTTG ATTATCCAAT CAAGCCTTTA ATTGAGCATG CTACTCCGTC ATTTAGGCAA
     Fla831cp-W CATATCTGGT GTCTGGGTTA TGATGGATGA TACTACAGGA ACCCAAGTTG ATTATCCAAT CAAGCCTTTA ATTGAGCATG CTACTCCGTC ATTTAGGCAA
                                                  INSERTION
```

FIG. 4C

```
                1001                                                1100
Australiancp-W ATTATGGCTC ACTTTAGTAA TGCGGCAGAA GCATATATTG CAAAGAGAAA TGCTACTGAG AGATACATGC CGCGGTATGG AATCAAGAGA AATTTGACTG
      Hacp-P   ATTATGGCTC ACTTTAGTAA CGCGGCAGAA GCATACATTG CGAAGAGAAA TGCTACTGAG AGGTACATGC CGCGGTATGG AATCAAGAGA AATTTGACTG
     Usacp-P   ATTATGGCTC ACTTTAGTAA CGCGGCAGAA GCATACATTG CGAAGAGAAA TGCTACTGAG AGGTACATGC CGCGGTATGG AATCAAGAGA AATTTGACTG
     Usacp-W   ATTATGGCTC ACTTTAGTAA CGCGGCAGAA GCATACATTG CGAAAAGAAA TGCTACTGAG AGGTACATGC CGCGGTATGG AATCAAGAGA AATTTGACTG
    Fla831cp-W ATTATGGCTC ACTTTAGTAA CGCGGCAGAA GCATACATTG CAAGGAGAAA TGCTACTGAG AGGTACATGC CGCGGTATGG AATCAAGAGA AATTTGACTG

                1101                                                1200
Australiancp-W ACATTAGCCT CGCCAGATAC GCTTTCGATT TCTATGAGGT GAATTCGAAA ACACCTGATA GGGCTCGCGA AGCTCACATG CAGATGAAAG CTGCAGCGCT
      Hacp-P   ACATTAGCCT CGCTAGATAC GCTTTCGACT TCTATGAGGT GAATTCGAAA ACACCTGATA GGGCTCGCGA AGCTCACATG CAGATGAAGG CTGCAGCGCT
     Usacp-P   ACATTAGCCT CGCTAGATAC GCTTTCGACT TCTATGAGGT GAATTCGAAA ACACCTGATA GGGCTCGCGA AGCTCACATG CAGATGAAGG CTGCAGCGCT
     Usacp-W   ACATTAGCCT CGCTAGATAC GCTTTCGACT TCTATGAGGT GAATTCGAAA ACACCTGATA GGGCTCGCGA AGCCCACATG CAGATGAAGG CTGCAGCACT
    Fla831cp-W ACATTAGCCT CGCTAGATAC GCTTTCGATT TCTATGAGGT TAATTCGAAA ACACCTGATA GGGCTCGCGA AGCTCGCATG CAGATGAAAG CTGCAGCGCT

                1201                                                1300
Australiancp-W GCGAAACACT AGTCGCAGAA TGTTTGGTAT GGACGGCAGT GTTAGTAACA AGGAAGAAAA CATGGAGAGA CACACAGTGG AAGATGTCAA TAGAGACATG
      Hacp-P   GCGAAACACC AGTCGCAGAA TGTTTCGTAT GGACGGCAGT GTTAGTAACA AGGAAGAAAA CACGGAGAGA CACACAGTGG AAGATGTCAA TAGAGACATG
     Usacp-P   GCGAAACACC AGTCGCAAAA TGTTTGGTAT GGACGGCAGT GTTAGTAACA AGGAAGAAAA CACGGAGAGA CACACAGTGG AAGATGTCAA TAGAGACATG
     Usacp-W   GCGAAACACT AGTCGCAGAA TGTTTGGTAT GGACGGCAGT GTTAGTAACA AGGAAGAAAA CACGGAGAGA CACACAGTGG AAGACGTCAA TAGAGACATG
    Fla831cp-W GCGAAACACT AATCGCAGAA TGTTTGGTAT GGACGGCAGT GTCAGTAACA AGGAAGAAAA TACGGAGAGA CACACAGTGG AAGATGTCAA TAGAGACATG

                1301                                                1400
Australiancp-W CACTCTCTCC TGGGTATGCG CAACTGAATA CTCGCACTTG TGTGTTTGTC GGGCCTGGCT CGACCTTGTT TCACCTTATA GTACTATATA AGCATTAGAA
      Hacp-P   CACTCTCTCC TGGGTATGCG CAACTAAATA CCTGCGCTTG TGTGTTTGTT GAGTCTGACT CGACCCTGTT TCACCTTATG GTACTATATA AGCATTAGAA
     Usacp-P   CACTCTCTCC TGGGTATGCG CAACTAAATA CCTGCGCTTG TGTGTTTGTT GAGTCTGACT CGACCCTGTT TCACCTTATG GTACTATATA AGCATTAGAA
     Usacp-W   CACTCTCTCC TGGGTATGCG CAACTAAATA CTTGCGCTTG TGTGTTTGTC GAGCTTGACT CGACCCTGTT TCACCTTATA GTACTATATA AGCATTAGAA
    Fla831cp-W CACTCTCTCC TGGGTATGCG CAACTGAATA CTCGCGCTTG TGTGTTTGTC GAGTCTAACT CGACCCTGTT TCACCCCATG G.......... ..........
```

FIG. 5A

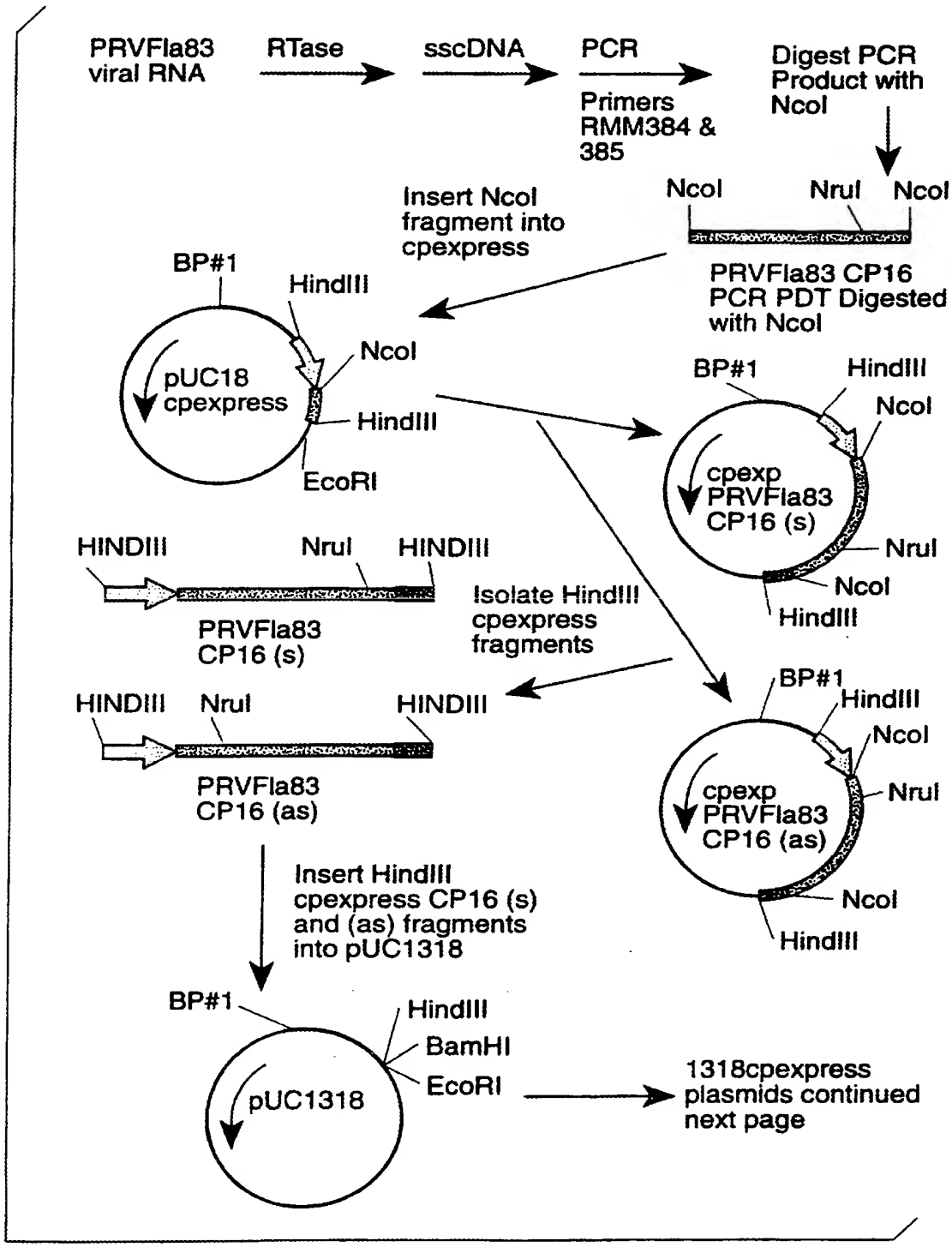
```
               1                                                                                                     100
        Hacp-P RNKQNLWFMS HRGILIDDIY IPKLEPERIV AILEWDKSKL PEHRLEAITA AMIESWGYGD LTHQIRRFYQ WVLEQAPFNE LAKQGRAPYV SEVGLRRLYT
       Usacp-P .......... .......... .......... .......... .......... .......... LTHQIRRFYQ WVLEQAPFNE LAKQGRAPYV SEVGLRRLYT
       Usacp-W .......... .......... .......... .......... .......ITA AMIESWGYGD LTHQIRRFYQ WVLEQAPFNE LAKQGRAPYV SEVGLRRLYT
    Fla831cp-W .......... .......... .......... .......... .......... .......... .......... .vleqAPFNE LAKQGRAPYV SEVGLRRLYT
Australiancp-W .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
                                                                                          LONG

               *101 *                    *            *  **        *     *                *             *    * 200
        Hacp-P SERGSMDELE AYIDKYFERE RGDSPELLVY HESRGTDDYQ LVCSNNTHVF HQSKNEAVDA GLNEKLKEKE KQKEKEKEKQ KEKEKDGASD GNDVSTSTKT
       Usacp-P SERGSMDELE AYIDKYFERE RGDSPELLVY HESRSTDDYQ LVCSNNTHVF HQSKNEAVDA GLNEKLKEKE NQKEKEKEKQ KEKEKDGASD GNDVSTSTKT
       Usacp-W SERGSMDELE AYIDKYFERE RGDSPELLVY HESRSTDDYQ LVCSNNTHVF HQSKNEAVDT GLNEKFKEKE KQKEKEKEKQ KEKEKDDASD GNDVSTSTKT
    Fla831cp-W CERGSVDELE AYIDKYFERE RGDSPEVLVY HESRSTDDYE LVRVNNTHVF HQAKNEAVDA GLNEKLKEKE KQREKEKEKQ KEKEKDDASD GNDVLTSTKT
Australiancp-W .......... .......... .......... .......... .......VF HQSKNEAVDA GLNEKLKEKE KQKEKEKEKQ KEKEKDDASD GNDVSTSTKT
               1   2                     3            4  56        7                                            8
                                                                 SHORT
               201            *  *        *        *                                 *          *  ***          300
        Hacp-P GERDRDVNVG TSGTFTVPRI KSFTDKMVLP RIKGKTVLNL NHLLQYNPQQ IDISNTRATH SQFEKWYEGV RNDYGLNDNE MQVMLNGLMV WCIENGTSPD
       Usacp-P GERDRDVNVG TSGTFTVPRI KSFTDKMVLP RIKGKTVLNL NHLLQYNPQQ IDISNTRATH SQFEKWYEGV RNDYGLNDNE MQVMLNGLMV WCIENGTSPD
       Usacp-W GERDRDVNVG TSGTFTVPRI KSFTDKMILP RIKGKSVLNL NHLLQYNPQQ IDISNTRATQ SQFEKWYEGV RNDYGLNDNE MQVMLNGLMV WCIENGTSPD
    Fla831cp-W GERDRDVNVG TSGTFTIPRI KPFNDKMILP RIKGKTVLNL NHLLQYNPQQ IDISNTRATQ SQFEKWHEGV RNDYGLNDKE MEVMLNGLMV WCIENGTSPD
Australiancp-W GERDRDVNVG TSGTFTVPRI KSFTDKMILP RIKGKTVLNL NHLLQYNPQQ IDISNTRATQ SQFEKWYEGV RNDYGLNDNE MQVMLNGLMV WCIENGTSPD
                                 9                                                10          11 12
```

FIG. 5B

```
                 301      * ***                 *               *                                                   400
         Hacp-P  ISGVWVMMDG ..ETQVDYPI KPLIEHATPS FRQIMAHFSN AAEAYIAKRN ATERYMPRYG IKRNLTDISL ARYAFDFYEV NSKTPDRARE AHMQMKAAAL
        Usacp-P  ISGVWVMMDG ..ETQVDYPI KPLIEHATPS FRQIMAHFSN AAEAYIAKRN ATERYMPRYG IKRNLTDISL ARYAFDFYEV NSKTPDRARE AHMQMKAAAL
        Usacp-W  ISGVWVMMDG ..ETQVDYPI KPLIEHATPS FRQIMAHFSN AAEAYIAKRN ATERYMPRYG IKRNLTDISL ARYAFDFYEV NSKTPDRARE AHMQMKAAAL
      Fla831cp-W ISGVWVMMDD TTGTQVDYPI KPLIEHATPS FRQIMAHFSN AAEAYIARRN ATERYMPRYG IKRNLTDISL ARYAFDFYEV NSKTPDRARE ARMQMKAAAL
  Australiancp-W ISGVWVMMDG ..ETQVDYPI KPLIEHATPT FRQIMAHFSN AAEAYIAKRN ATERYMPRYG IKRNLTDISL ARYAFDFYEV NSKTPDRARE AHMQMKAAAL
                        13 INSERTION                        14

                 401*             *
         Hacp-P  RNTSRRMFGM DGSVSNKEEN TERHTVEDVN RDMHSLLGMR N*
        Usacp-P  RNTSRKMFGM DGSVSNKEEN TERHTVEDVN RDMHSLLGMR N*
        Usacp-W  RNTSRRMFGM DGSVSNKEEN TERHTVEDVN RDMHSLLGMR N*
      Fla831cp-W RNTNRRMFGM DGSVSNKEEN TERHTVEDVN RDMHSLLGMR N*
  Australiancp-W RNTSRRMFGM DGSVSNKEEN MERHTVEDVN RDMHSLLGMR N*
```

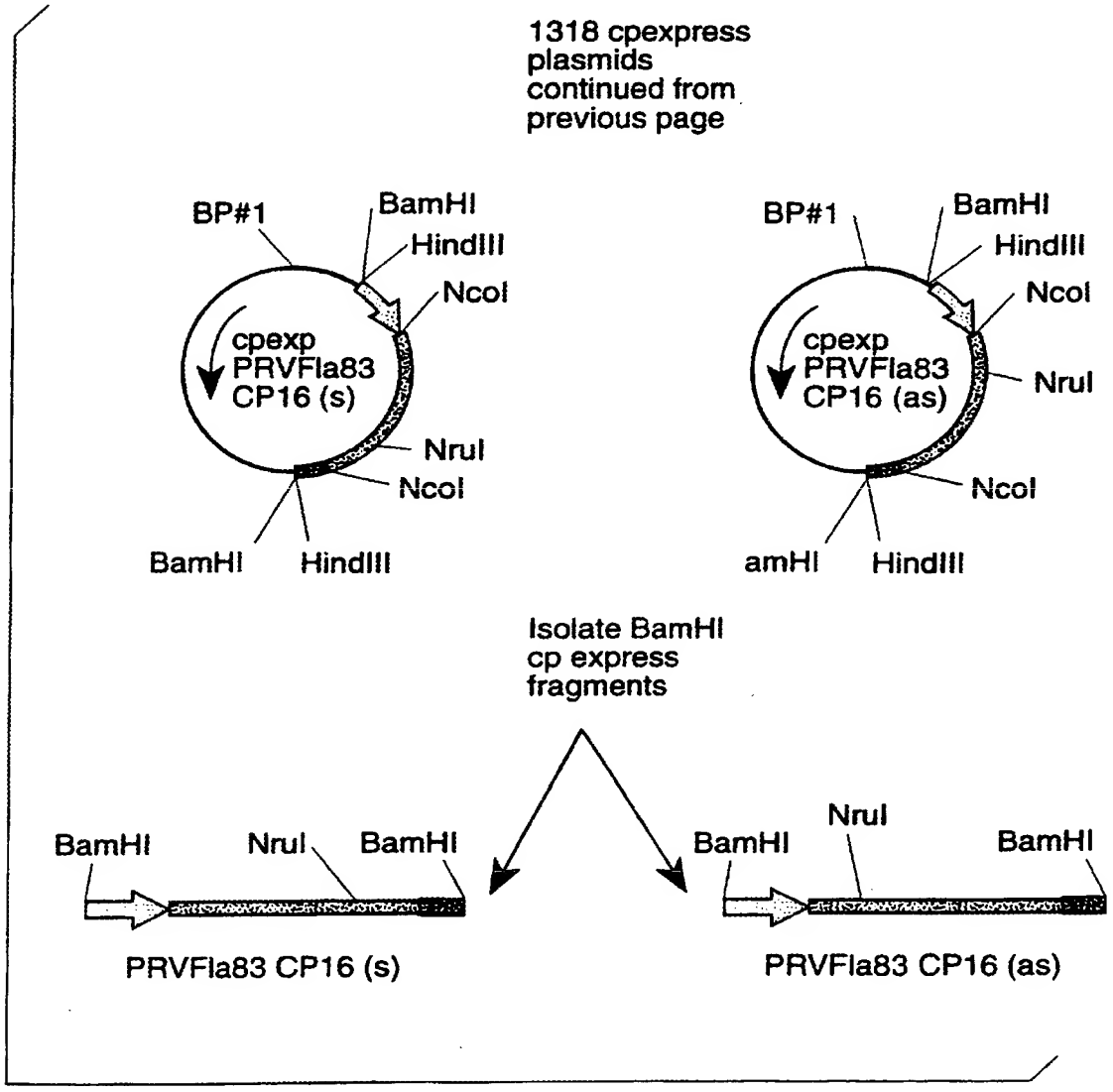
13/22

FIG. 6A


PRVFla83 viral RNA
RTase
sscDNA
PCR
Primers RMM384 & 385
Digest PCR Product with NcoI
NcoI
NruI
NcoI
Insert NcoI fragment into cpexpress
PRVFla83 CP16 PCR PDT Digested with NcoI
BP#1
HindIII
NcoI
pUC18 cpexpress
HindIII
EcoRI
BP#1
HindIII
NcoI
cpexp PRVFla83 CP16 (s)
NruI
NcoI
HindIII
HINDIII
NruI
HINDIII
PRVFla83 CP16 (s)
Isolate HindIII cpexpress fragments
HINDIII
NruI
HINDIII
PRVFla83 CP16 (as)
BP#1
HindIII
NcoI
cpexp PRVFla83 CP16 (as)
NruI
NcoI
HindIII
Insert HindIII cpexpress CP16 (s) and (as) fragments into pUC1318
BP#1
HindIII
BamHI
EcoRI
pUC1318
1318cpexpress plasmids continued next page

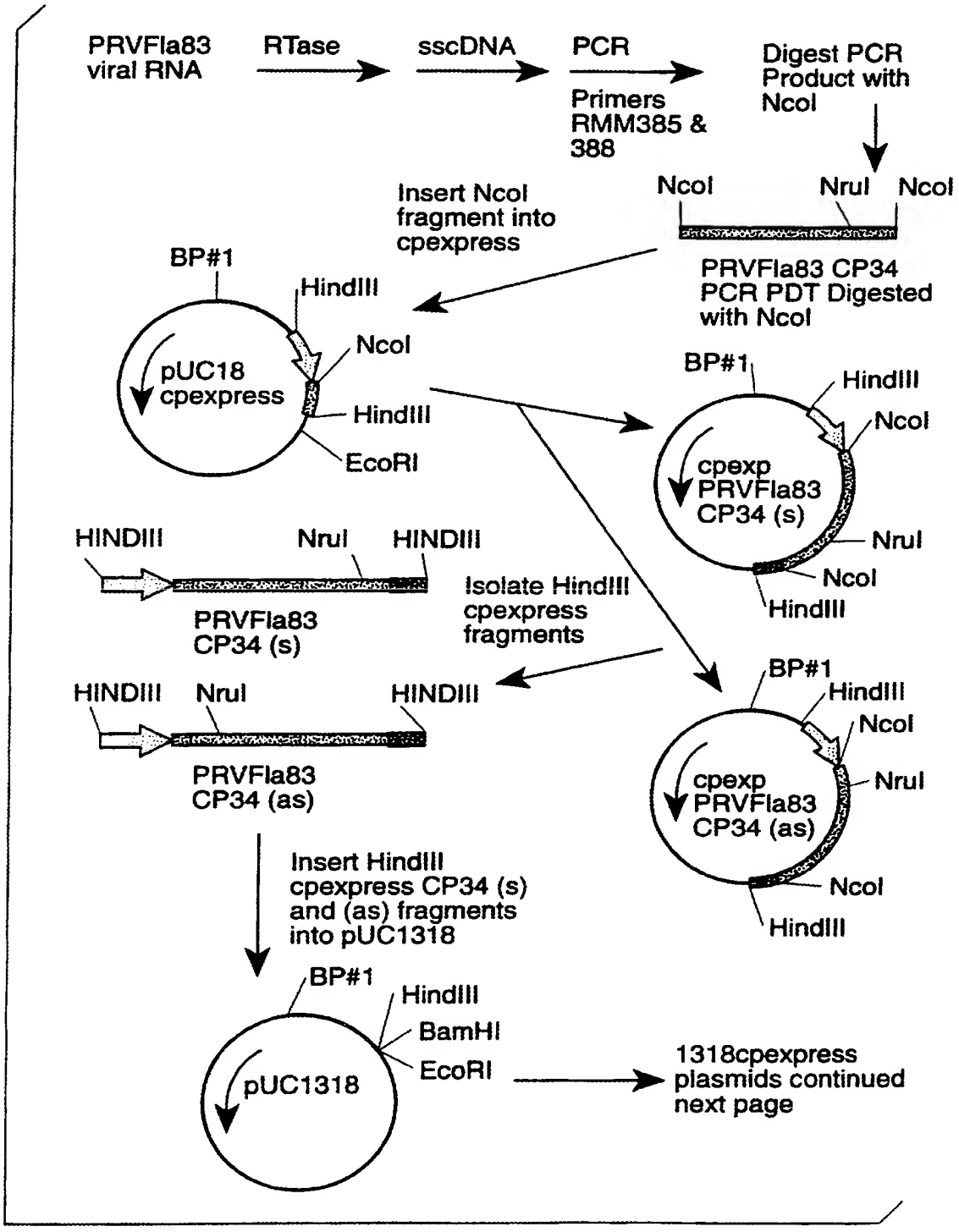
14/22

FIG. 6B


1318 cpexpress plasmids continued from previous page
BP#1
BamHI
HindIII
NcoI
cpexp PRVFla83 CP16 (s)
NruI
NcoI
BamHI
HindIII
BP#1
BamHI
HindIII
NcoI
cpexp PRVFla83 CP16 (as)
NruI
NcoI
amHI
HindIII
Isolate BamHI cp express fragments
BamHI
NruI
BamHI
PRVFla83 CP16 (s)
BamHI
NruI
BamHI
PRVFla83 CP16 (as)

15/22

FIG. 7A


PRVFla83 viral RNA
RTase
sscDNA
PCR
Primers RMM385 & 388
Digest PCR Product with NcoI
NcoI
NruI
NcoI
PRVFla83 CP34 PCR PDT Digested with NcoI
Insert NcoI fragment into cpexpress
BP#1
HindIII
NcoI
pUC18 cpexpress
HindIII
EcoRI
BP#1
HindIII
NcoI
cpexp PRVFla83 CP34 (s)
NruI
NcoI
HindIII
HINDIII
NruI
HINDIII
PRVFla83 CP34 (s)
Isolate HindIII cpexpress fragments
HINDIII
NruI
HINDIII
PRVFla83 CP34 (as)
BP#1
HindIII
NcoI
cpexp PRVFla83 CP34 (as)
NruI
NcoI
HindIII
Insert HindIII cpexpress CP34 (s) and (as) fragments into pUC1318
BP#1
HindIII
BamHI
EcoRI
pUC1318
1318cpexpress plasmids continued next page

16/22

FIG. 7B

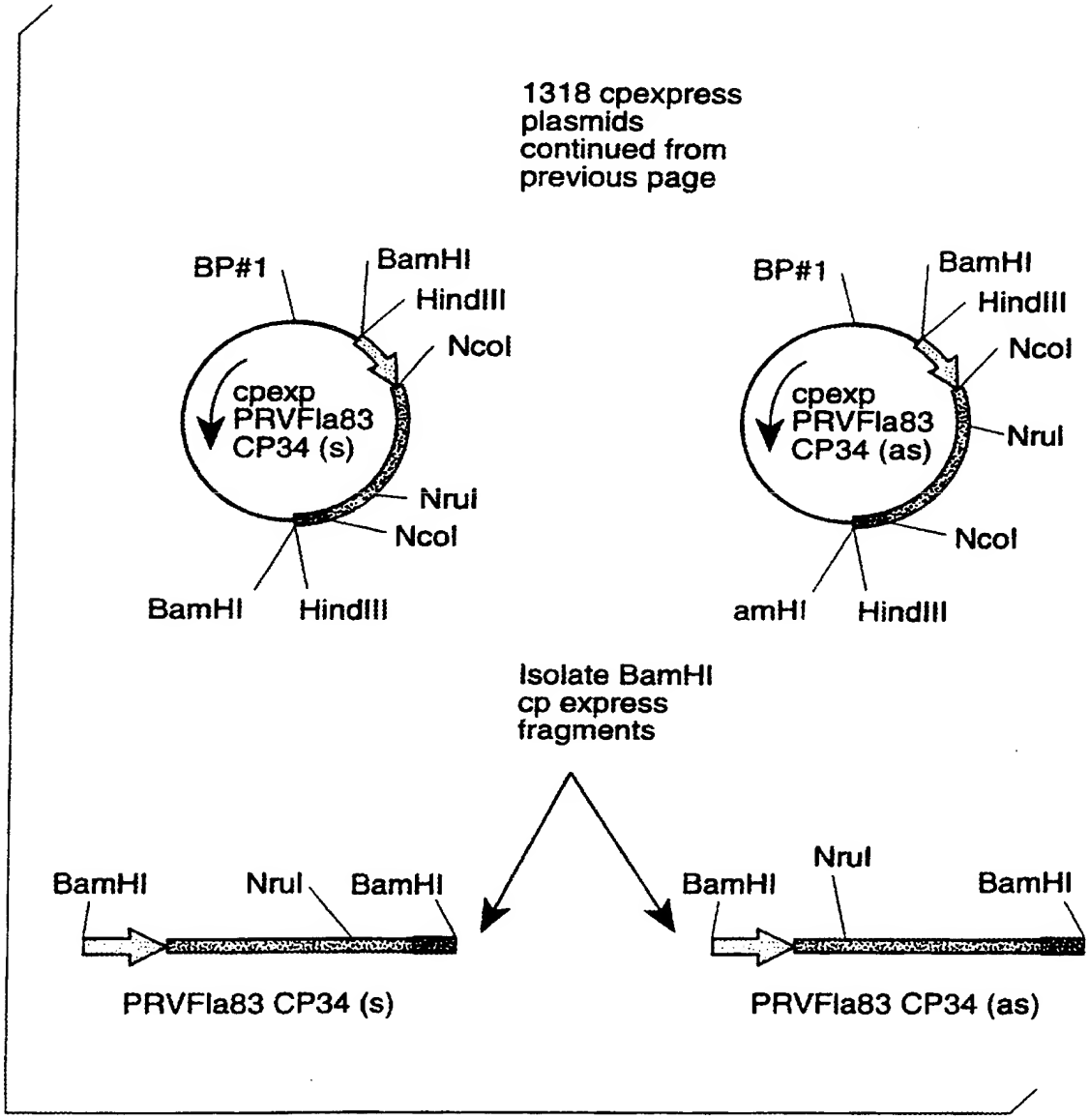

1318 cpexpress plasmids continued from previous page
BP#1
BamHI
HindIII
NcoI
cpexp PRVFla83 CP34 (s)
NruI
NcoI
BamHI
HindIII
BP#1
BamHI
HindIII
NcoI
cpexp PRVFla83 CP34 (as)
NruI
NcoI
amHI
HindIII
Isolate BamHI cp express fragments
BamHI
NruI
BamHI
PRVFla83 CP34 (s)
BamHI
NruI
BamHI
PRVFla83 CP34 (as)

FIG. 8A

```
                  201                                                                                                      300
   Fla831cpW  .......... .......... .CCATGGCTC CATTCAATGA GCTGGCGAAA CAAGGGAGGG CCCCATACGT CTCGGAAGTT GGATTAAGAA GGTTGTATAC
       HacpP  GATTTTACCA ATGGGTTCTT GAGCAAGCTC CATTCAATGA GTTGGCGAAA CAAGGAAGGG CCCCATACGT CTCGGAAGTT GGATTAAGAA GATTGTACAC
      UsacpP  GATTTTACCA ATGGGTTCTT GAGCAAGCTC CATTCAATGA GTTGGCGAAA CAAGGAAGGG CCCCATACGT CTCGGAAGTT GGATTAAGAA GATTGTACAC
      UsacpW  GATTTTACCA ATGGGTTCTT GAGCAAGCTC CATTCAATGA GTTGGCGAAA CAAGGAAGGG CCCCATACGT CTCGGAAGTT GGATTAAGAA GATTGTACAC
AustraliancpW  .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
    BrazilcpP  .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
       YkcpP  .......... .......... .agcaggctc cattcaatga attagcgaag cagggcaggg ctccatatgt gtctgaggtt ggattgaggc gcttatacac

                  301                                                                                                      400
   Fla831cpW  GTGTGAACGC GGATCAGTGG ATGAATTGGA AGCGTATATA GATAAATATT TTGAGCGTGA GAGGGGAGAC TCACCCGAAG TACTGGTGTA CCATGAATCA
       HacpP  AAGTGAACGT GGATCAATGG ACGAATTAGA AGCGTATATA GATAAATACT TTGAGCGTGA GAGAGGAGAC TCGCCCGAAT TACTAGTGTA CCATGAATCA
      UsacpP  AAGTGAACGT GGATCAATGG ACGAATTAGA AGCGTATATA GATAAATACT TTGAGCGTGA GAGAGGAGAC TCGCCCGAAT TACTAGTGTA CCATGAATCA
      UsacpW  AAGTGAACGT GGATCAATGG ATGAATTAGA AGCGTATATA GATAAATACT TTGAGCGTGA GAGAGGAGAC TCACCCGAAT TACTAGTGTA CCATGAATCA
AustraliancpW  .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
    BrazilcpP  .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
       YkcpP  tagtaagcgt gggtcaatgg atgaattgga ggcctacata gataaatatt ttgaacgaga gagaggagac tcacccgaat tactagtgta ccatgaatcg

                  401                                                                                                      500
   Fla831cpW  AGGAGTACTG ATGATTATGA ACTTGTTCGT GTCAACAATA CACATGTGTT TCATCAAGCC AAAAATGAAG CTGTGGACGC TGGTTTGAAC GAAAAGCTCA
       HacpP  AGGGGCACTG ATGATTATCA ACTTGTTTGT AGCAACAATA CGCATGTGTT TCATCAGTCC AAGAATGAAG CTGTGGATGC TGGTTTGAAT GAAAAACTCA
      UsacpP  AGGAGCACTG ATGATTATCA ACTTGTTTGT AGCAACAATA CGCATGTGTT TCATCAGTCC AAGAATGAAG CTGTGGATGC TGGTTTGAAT GAAAAACTCA
      UsacpW  AGGAGCACTG ATGATTATCA ACTTGTTTGC AGTAACAATA CACATGTGTT TCATCAGTCC AAAAATGAAG CTGTGGATAC TGGTTTGAAT GAAAAATTCA
AustraliancpW  .......... .......... .......... .......... .......... .......TCC AAAAATGAAG CTGTGGATGC TGGTTTGAAC GAAAAGCTCA
    BrazilcpP  .......... .......... .......... .......... .......... .......... ....CCATGG CTGTGGATGC TGGTTTGAAT GAAAAGCGTA
       YkcpP  agaagcactg acaatcatca attaacccgc ggcagtaata cacatgtgtt tcaccagtct aaaaatgaag ctgtggatac cggtctgaat gagaagctca

                  501                                                                                                      600
   Fla831cpW  AAGAAAAAGA AAAACAGAGA GAGAAAGAAA AAGAAAAACA AAAAGAGAAA GAAAAAGATG ATGCTAGTGA CGGAAATGAT GTGTTAACTA GCACAAAAAC
       HacpP  AAGAGAAGGA AAAACAGAAA GAAAAAGAAA AAGAAAAACA AAAAGAGAAA GAAAAAGACG GTGCTAGTGA CGGAAATGAT GTGTCAACTA GCACAAAAAC
      UsacpP  AAGAGAAGGA AAATCAGAAA GAAAAAGAAA AAGAAAAACA AAAAGAGAAA GAAAAAGACG GTGCTAGTGA CGGAAATGAT GTGTCAACTA GCACAAAAAC
      UsacpW  AAGAAAAGGA AAAACAGAAA GAAAAAGAAA AAGAAAAACA AAAAGAGAAA GAAAAAGACG ATGCTAGTGA CGGAAATGAT GTGTCAACTA GCACAAAAAC
AustraliancpW  AAGAAAAAGA AAAACAGAAA GAAAAAGAAA AAGAAAAACA AAAAGAGAAA GAGAAAGACG ATGCTAGTGA CGGAAATGAT GTGTCAACTA GCACAAAAAC
    BrazilcpP  AAGAACAAGA GAAACAAGAA GAAAAAG... AAGAAAAACA AAAAAAGAAA GAAAAAGACG ATGCTAGTTA CGGAAACGAT GTGTCAACTA GCACAAGAAC
       YkcpP  aagaaaaaga aaagcagaaa gaaaaagaaa aagataaaca acaagataaa gacaatgatg gagctagtga cggaaacgat gtgtcaacta gcacaaaaac
```

FIG. 8B

```
             601                                                                                                         700
Fla83lcpW    TGGAGAGAGA GATAGAGATG TCAATGTTGG AACTAGTGGG ACTTTCACTA TTCCAAGGAT TAAACCATTC AATGATAAGA TGATTTTACC GAGAATTAAG
HacpP        TGGAGAGAGA GATAGAGATG TCAATGTTGG GACCAGTGGA ACTTTCACTG TTCCGAGAAT TAAATCATTT ACTGATAAGA TGGTTCTACC GAGAATTAAG
UsacpP       TGGAGAGAGA GATAGAGATG TCAATGTTGG GACCAGTGGA ACTTTCACTG TTCCGAGAAT TAAATCATTT ACTGATAAGA TGGTTCTACC GAGAATTAAG
UsacpW       TGGAGAGAGA GATAGAGATG TCAATGTTGG GACCAGTGGA ACTTTCACTG TTCCGAGAAT TAAATCATTT ACTGATAAGA TGATTCTACC GAGAATTAAG
AustraliancpW TGGAGAGAGA GATAGAGATG TCAATGTTGG GACCAGTGGA ACTTTCACTG TTCCAAGAAT CAAATCATTT ACTGACAAGA TGATTCTACC AAGAATTAAG
BrazilcpP    TGGAGAGAGA GACAGAGATG TCAATGTTGG GACCAGTGGA ACTTTCACTG TTCCGAGAAC AAAATCATTT ACTGATAAGA TGATTTTACC TAGAATTAAG
YkcpP        tggagagaga gatagggatg tcaatgccgg aactagtgga accttcactg ttccgaggat aaagtcattt actgataaga tgatcttacc aagaattaag

             701                                                                                                         800
Fla83lcpW    GGAAAAACTG TCCTTAATTT AAATCATCTT CTTCAGTATA ATCCGCAACA AATTGACATT TCGAACACTC GTGCCACTCA GTCACAATTT GAAAAATGGC
HacpP        GGGAAGACTG TCCTTAATTT AAATCATCTT CTTCAGTACA ATCCGCAACA AATTGACATT TCTAACACTC GTGCCACTCA TTCACAATTT GAGAAGTGGT
UsacpP       GGGAAGACTG TCCTTAATTT AAATCATCTT CTTCAGTACA ATCCGCAACA AATTGACATT TCTAACACTC GTGCCACTCA TTCACAATTT GAGAAGTGGT
UsacpW       GGAAAGTCTG TCCTTAATTT AAATCACCTA CTTCAGTATA ATCCGCAACA AATTGACATT TCTAACACTC GTGCCACTCA GTCACAATTT GAGAAGTGGT
AustraliancpW GGAAAGACTG TCCTTAATTT AAATCACCTT CTTCAGTATA ACCCGCAACA AATTGACATT TCTAACACTC GTGCCACTCA GTCACAATTT GAGAAGTGGT
BrazilcpP    GGAAAAACTG TCCTTAATTT AAATCATCTG ATTCAGTATA ATCCGCAACA AATTGACATT TCTAACACTC GTGCTACTCA ATCACAATTT GAGAAGTGGT
YkcpP        ggaaaaactg tccttaattt aaatcatctt cttcagtata atccgaaaca agttgacatc tcaaacactc gcgccactca atctcaattt gagaagtggt

             801                                                                                                         900
Fla83lcpW    ACGAGGGAGT GAGGAATGAT TATGGCCTGA ATGATAAAGA GATGGAAGTA ATGTTAAATG GCTTGATGGT TTGGTGTATT GAGAATGGTA CATCTCCGGA
HacpP        ATGAGGGAGT GAGGAATGAT TATGGCCTTA ATGATAATGA AATGCAAGTG ATGCTAAATG GTTTGATGGT TTGGTGTATC GAGAATGGTA CATCTCCAGA
UsacpP       ATGAGGGAGT GAGGAATGAT TATGGCCTTA ATGATAATGA AATGCAAGTG ATGCTAAATG GTTTGATGGT TTGGTGTATC GAGAATGGTA CATCTCCAGA
UsacpW       ATGAGGGAGT GAGGAATGAT TATGGCCTTA ATGATAATGA AATGCAAGTG ATGCTAAATG GTTTGATGGT TTGGTGTATC GAGAATGGTA CATCTCCAGA
AustraliancpW ATGAGGGAGT GAGGAATGAT TATGGCCTTA ATGATAATGA AATGCAAGTG ATGCTAAATG GCTTGATGGT TTGGTGTATC GAGAATGGTA CATCTCCAGA
BrazilcpP    ACGAGGGAGT GAGGAATGAT TATGGCCTTA ATGATAATGA GATGCAAATA GTGCTAAATG GTTTGATGGT TTGGTGTATC GAAAACGGTA CATCTCCAGA
YkcpP        atgagggagt gagaaatgat tatggcctta atgataacga aatgcaagta atgttaaatg gtttgatggt ttggtgtatc gaaaatggta catctccaga

             901                                                                                                        1000
Fla83lcpW    CATATCTGGT GTCTGGGTTA TGATGGATGA TACTACAGGA ACCCAAGTTG ATTATCCAAT CAAGCCTTTA ATTGAGCATG CTACTCCGTC ATTTAGGCAA
HacpP        CATATCTGGT GTCTGGGTTA TGATGGATG. .....GGGAA ACCCAAGTTG ATTATCCAAT CAAGCCTTTG ATTGAGCATG CTACTCCGTC ATTTAGGCAA
UsacpP       CATATCTGGT GTCTGGGTTA TGATGGATG. .....GGGAA ACCCAAGTTG ATTATCCAAT CAAGCCTTTG ATTGAGCATG CTACTCCGTC ATTTAGGCAA
UsacpW       CATATCTGGT GTCTGGGTTA TGATGGATG. .....GGGAA ACCCAAGTTG ATTATCCAAT CAAGCCTTTA ATTGAGCATG CTACTCCGTC ATTTAGGCAA
AustraliancpW CATATCTGGT GTCTGGGTTA TGATGGATG. .....GGGAA ACCCAAGTTG ATTATCCAAT CAAGCCTTTA ATTGAGCATG CTACTCCGAC ATTTAGGCAA
BrazilcpP    CATATCTGGT GTCTGGGTTA TGATGGATG. .....GGGAA ACCCAGGTTG ACTATCCAAT CAAGCCTTTA ATTGAGCATG CTACTCCGTC GTTTAGGCAA
YkcpP        tatatctggt gtctgggtta tgatggatg. .....gggaa acccaagtcg attatcccat taaacctttg attgaacacg caactccttc atttaggcaa
```

FIG. 8C

```
               1001                                                                                                          1100
Fla831cpW      ATTATGGCTC ACTTTAGTAA CGCGGCAGAA GCATACATTG CAAGGAGAAA TGCTACTGAG AGGTACATGC CGCGGTATGG AATCAAGAGA AATTTGACTG
HacpP          ATTATGGCTC ACTTTAGTAA CGCGGCAGAA GCATACATTG CGAAGAGAAA TGCTACTGAG AGGTACATGC CGCGGTATGG AATCAAGAGA AATTTGACTG
UsacpP         ATTATGGCTC ACTTTAGTAA CGCGGCAGAA GCATACATTG CGAAGAGAAA TGCTACTGAG AGGTACATGC CGCGGTATGG AATCAAGAGA AATTTGACTG
UsacpW         ATTATGGCTC ACTTTAGTAA CGCGGCAGAA GCATACATTG CGAAAAGAAA TGCTACTGAG AGGTACATGC CGCGGTATGG AATCAAGAGA AATTTGACTG
AustraliancpW  ATTATGGCTC ACTTTAGTAA TGCGGCAGAA GCATATATTG CAAAGAGAAA TGCTACTGAG AGATACATGC CGCGGTATGG AATCAAGAGA AATTTGACTG
BrazilcpP      ATTATGGCTC ATTTCAGTAA CGCGGCAGAA GCATACATTA CAAAGAGAAA TGCTACTGAG AGGTACATGC CGCGGTATGG GATCAAGAGA AATTTGACTG
YkcpP          atcatggctc acttcagtaa cgcggcagag gcatacatcg cgaagaggaa tgcaactgag aagtacatgc cgcggtatgg aatcaagaga aatttgactg

               1101                                                                                                          1200
Fla831cpW      ACATTAGCCT CGCTAGATAC GCTTTCGATT TCTATGAGGT TAATTCGAAA ACACCTGATA GGGCTCGCGA AGCTCGCATG CAGATGAAAG CTGCAGCGCT
HacpP          ACATTAGCCT CGCTAGATAC GCTTTCGACT TCTATGAGGT GAATTCGAAA ACACCTGATA GGGCTCGCGA AGCTCACATG CAGATGAAGG CTGCAGCGCT
UsacpP         ACATTAGCCT CGCTAGATAC GCTTTCGACT TCTATGAGGT GAATTCGAAA ACACCTGATA GGGCTCGCGA AGCTCACATG CAGATGAAGG CTGCAGCGCT
UsacpW         ACATTAGCCT CGCTAGATAC GCTTTCGACT TCTATGAGGT GAATTCGAAA ACACCTGATA GGGCTCGCGA AGCCCACATG CAGATGAAGG CTGCAGCACT
AustraliancpW  ACATTAGCCT CGCCAGATAC GCTTTCGATT TCTATGAGGT GAATTCGAAA ACACCTGATA GGGCTCGCGA AGCTCACATG CAGATGAAAG CTGCAGCGCT
BrazilcpP      ACATTAGTCT TGCTAGATAT GCTTTCGATT TCTATGAGGT GAATTCGAAA ACACCTGATA GGGCTCGCGA AGCTCACATG CAGATGAAAG CTGCAGCGCT
YkcpP          acattagtct cgctagatat gctttcgatt tctatgaggt gaattcgaaa acacctgata gggctcgtga agctcatatg cagatgaagg ctgcagcgct

               1201                                                                                                          1300
Fla831cpW      GCGAAACACT AATCGCAGAA TGTTTGGTAT GGACGGCAGT GTCAGTAACA AGGAAGAAAA TACGGAGAGA CACACAGTGG AAGATGTCAA TAGAGACATG
HacpP          GCGAAACACC AGTCGCAGAA TGTTTGGTAT GGACGGCAGT GTTAGTAACA AGGAAGAAAA CACGGAGAGA CACACAGTGG AAGATGTCAA TAGAGACATG
UsacpP         GCGAAACACC AGTCGCAAAA TGTTTGGTAT GGACGGCAGT GTTAGTAACA AGGAAGAAAA CACGGAGAGA CACACAGTGG AAGATGTCAA TAGAGACATG
UsacpW         GCGAAACACT AGTCGCAGAA TGTTTGGTAT GGACGGCAGT GTTAGTAACA AGGAAGAAAA CACGGAGAGA CACACAGTGG AAGACGTCAA TAGAGACATG
AustraliancpW  GCGAAACACT AGTCGCAGAA TGTTTGGTAT GGACGGCAGT GTTAGTAACA AGGAAGAAAA CATGGAGAGA CACACAGTGG AAGATGTCAA TAGAGACATG
BrazilcpP      GCGAAACACT AATCGCAGAA TGTTTGGTAT GGACGGCAGT GTTAGTAACA AGGAAGAAAA CACGGAGAGA CACACAGTGG AAGATGTCAA TAGAGACATG
YkcpP          acgcaatact aatcgcaaaa tgtttggaat ggacggcagt gtcagtaaca aggaagaaaa cacggagaga cacacagtgg aagatgtcaa cagagacatg

               1301
Fla831cpW      CACTCTCTCC TGGGTATGCG CAACTGAATA CTCGCGCTTG TGTGTTTGTC GAGTCTAACT CGACCCTGTT TCACCCCATG G
HacpP          CACTCTCTCC TGGGTATGCG CAACTAAATA CCTGCGCTTG TGTGTTTGTT GAGTCTGACT CGACCCTGTT TCACCTTATG G
UsacpP         CACTCTCTCC TGGGTATGCG CAACTAAATA CCTGCGCTTG TGTGTTTGTT GAGTCTGACT CGACCCTGTT TCACCTTATG G
UsacpW         CACTCTCTCC TGGGTATGCG CAACTAAATA CTTGCGCTTG TGTGTTTGTC GAGCTTGACT CGACCCTGTT TCACCTTATA G
AustraliancpW  CACTCTCTCC TGGGTATGCG CAACTGAATA CTCGCACTTG TGTGTTTGTC GGGCCTGGCT CGACCTTGTT TCACCTTATA G
BrazilcpP      CACTCTCTCC TGGGTATGCG CAACTGAATA CTCGCGCT.. GGTGCTTGTT GGGCCTGACT CGACCCTGTT TCACCCCATG G
YkcpP          cactctctcc tgggtatgcg caattgaata ctcgcgctag tgtgtttgtc gggcctggct cgaccctgtt tcaccttata a
```

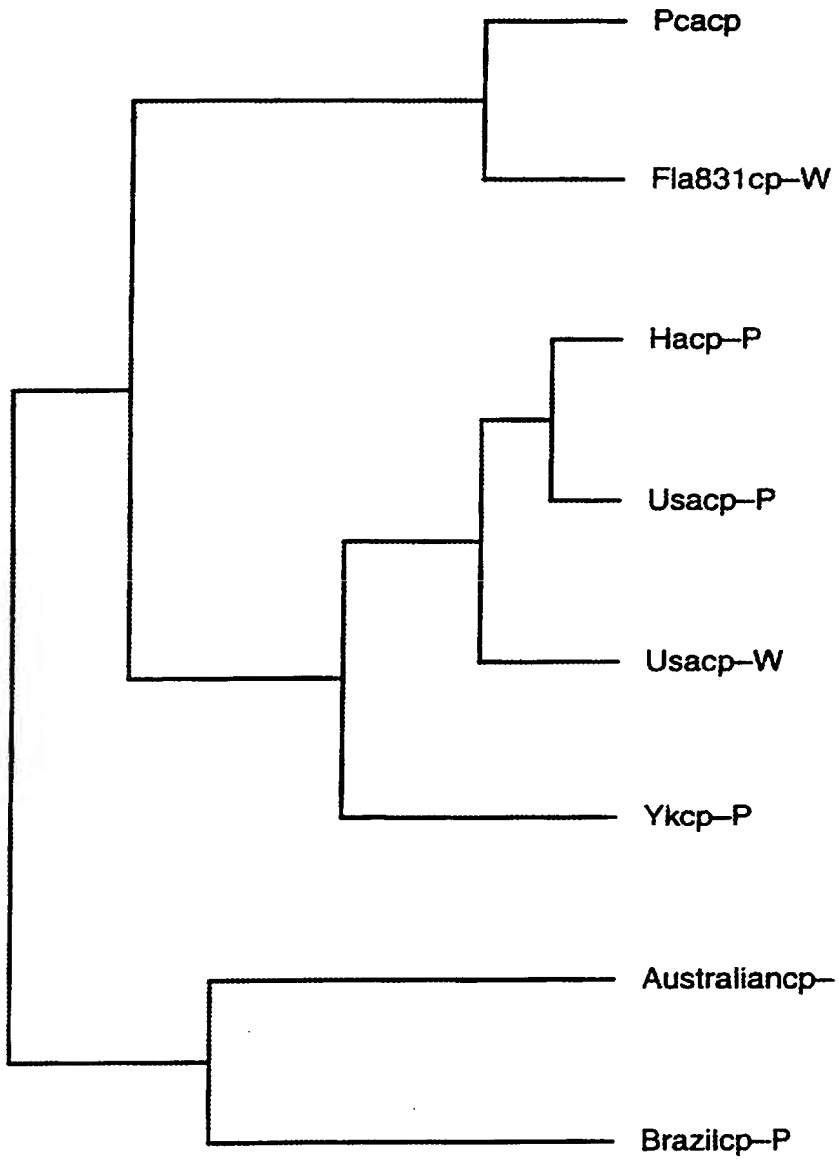
FIG. 9A

```
             1                                                                                                          100
Fla831cpW    .......... .......... .......... .......... .......... .......... .......... ....MAPFNE LAKQGRAPYV SEVGLRRLYT
HacpP        RNKQNLWFMS HRGILIDDIY IPKLEPERIV AILEWDKSKL PEHRLEAITA AMIESWGYGD LTHQIRRFYQ WVLEQAPFNE LAKQGRAPYV SEVGLRRLYT
UsacpP       .......... .......... .......... .......... .......... .......... LTHQIRRFYQ WVLEQAPFNE LAKQGRAPYV SEVGLRRLYT
UsacpW       .......... .......... .......... .......... .......ITA AMIESWGYGD LTHQIRRFYQ WVLEQAPFNE LAKQGRAPYV SEVGLRRLYT
YkcpP        .......... .......... .......... .......... .......... .......... .......... ...EQAPFNE LAKQGRAPYV SEVGLRRLYT
AustraliancpW .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
BrazilcpP    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........

             101                                                                                                        200
Fla831cpW    CERGSVDELE AYIDKYFERE RGDSPEVLVY HESRSTDDYE LVRVNNTHVF HQAKNEAVDA GLNEKLKEKE KQREKEKEKQ KEKEKDDASD GNDVLTSTKT
HacpP        SERGSMDELE AYIDKYFERE RGDSPELLVY HESRGTDDYQ LVCSNNTHVF HQSKNEAVDA GLNEKLKEKE KQKEKEKEKQ KEKEKDGASD GNDVSTSTKT
UsacpP       SERGSMDELE AYIDKYFERE RGDSPELLVY HESRSTDDYQ LVCSNNTHVF HQSKNEAVDA GLNEKLKEKE NQKEKEKEKQ KEKEKDGASD GNDVSTSTKT
UsacpW       SERGSMDELE AYIDKYFERE RGDSPELLVY HESRSTDDYQ LVCSNNTHVF HQSKNEAVDT GLNEKFKEKE KQKEKEKEKQ KEKEKDDASD GNDVSTSTKT
YkcpP        SKRGSMDELE AYIDKYFERE RGDSPELLVY HESRSTDNHQ LTRGSNTHVF HQSKNEAVDT GLNEKLKEKE KQKEKEKDKQ QDKDNDGASD GNDVSTSTKT
AustraliancpW .......... .......... .......... .......... .......... ..SKNEAVDA GLNEKLKEKE KQKEKEKEKQ KEKEKDDASD GNDVSTSTKT
BrazilcpP    .......... .......... .......... .......... .......... .....MAVDA GLNEKRKEQE KQEEKE_EKQ KKKEKDDASY GNDVSTSTRT

             201                                                                                                        300
Fla831cpW    GERDRDVNVG TSGTFTIPRI KPFNDKMILP RIKGKTVLNL NHLLQYNPQQ IDISNTRATQ SQFEKWHEGV RNDYGLNDKE MEVMLNGLMV WCIENGTSPD
HacpP        GERDRDVNVG TSGTFTVPRI KSFTDKMVLP RIKGKTVLNL NHLLQYNPQQ IDISNTRATH SQFEKWYEGV RNDYGLNDNE MQVMLNGLMV WCIENGTSPD
UsacpP       GERDRDVNVG TSGTFTVPRI KSFTDKMVLP RIKGKTVLNL NHLLQYNPQQ IDISNTRATH SQFEKWYEGV RNDYGLNDNE MQVMLNGLMV WCIENGTSPD
UsacpW       GERDRDVNVG TSGTFTVPRI KSFTDKMILP RIKGKSVLNL NHLLQYNPQQ IDISNTRATQ SQFEKWYEGV RNDYGLNDNE MQVMLNGLMV WCIENGTSPD
YkcpP        GERDRDVNAG TSGTFTVPRI KSFTDKMILP RIKGKTVLNL NHLLQYNPKQ VDISNTRATQ SQFEKWYEGV RNDYGLNDNE MQVMLNGLMV WCIENGTSPD
AustraliancpW GERDRDVNVG TSGTFTVPRI KSFTDKMILP RIKGKTVLNL NHLLQYNPQQ IDISNTRATQ SQFEKWYEGV RNDYGLNDNE MQVMLNGLMV WCIENGTSPD
BrazilcpP    GERDRDVNVG TSGTFTVPRT KSFTDKMILP RIKGKTVLNL NHLIQYNPQQ IDISNTRATQ SQFEKWYEGV RNDYGLNDNE MQIVLNGLMV WCIENGTSPD
```

FIG. 9B

```
                301                                                                                                          400
     Fla831cpW  ISGVWVMMDD TTGTQVDYPI KPLIEHATPS FRQIMAHFSN AAEAYIARRN ATERYMPRYG IKRNLTDISL ARYAFDFYEV NSKTPDRARE ARMQMKAAAL
         HacpP  ISGVWVMMDG ..ETQVDYPI KPLIEHATPS FRQIMAHFSN AAEAYIAKRN ATERYMPRYG IKRNLTDISL ARYAFDFYEV NSKTPDRARE AHMQMKAAAL
        UsacpP  ISGVWVMMDG ..ETQVDYPI KPLIEHATPS FRQIMAHFSN AAEAYIAKRN ATERYMPRYG IKRNLTDISL ARYAFDFYEV NSKTPDRARE AHMQMKAAAL
        UsacpW  ISGVWVMMDG ..ETQVDYPI KPLIEHATPS FRQIMAHFSN AAEAYIAKRN ATERYMPRYG IKRNLTDISL ARYAFDFYEV NSKTPDRARE AHMQMKAAAL
         YkcpP  ISGVWVMMDG ..ETQVDYPI KPLIEHATPS FRQIMAHFSN AAEAYIAKRN ATEKYMPRYG IKRNLTDISL ARYAFDFYEV NSKTPDRARE AHMQMKAAAL
AustraliancpW  ISGVWVMMDG ..ETQVDYPI KPLIEHATPT FRQIMAHFSN AAEAYIAKRN ATERYMPRYG IKRNLTDISL ARYAFDFYEV NSKTPDRARE AHMQMKAAAL
      BrazilcpP  ISGVWVMMDG ..ETQVDYPI KPLIEHATPS FRQIMAHFSN AAEAYITKRN ATERYMPRYG IKRNLTDISL ARYAFDFYEV NSKTPDRARE AHMQMKAAAL

                401                                                                                                          500
     Fla831cpW  RNTNRRMFGM DGSVSNKEEN TERHTVEDVN RDMHSLLGMR N*
         HacpP  RNTSRRMFGM DGSVSNKEEN TERHTVEDVN RDMHSLLGMR N*
        UsacpP  RNTSRKMFGM DGSVSNKEEN TERHTVEDVN RDMHSLLGMR N*
        UsacpW  RNTSRRMFGM DGSVSNKEEN TERHTVEDVN RDMHSLLGMR N*
         YkcpP  RNTNRKMFGM DGSVSNKEEN TERHTVEDVN RDMHSLLGMR N*
AustraliancpW  RNTSRRMFGM DGSVSNKEEN MERHTVEDVN RDMHSLLGMR N*
      BrazilcpP  RNTNRRMFGM DGSVSNKEEN TERHTVEDVN RDMHSLLGMR N*
```

FIG. 10


Pcacp
Fla831cp–W
Hacp–P
Usacp–P
Usacp–W
Ykcp–P
Australiancp–
Brazilcp–P

INTERNATIONAL SEARCH REPORT

Internation. pplication No
PCT/US 95/07272

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/40 C12N15/82 C12N1/21 C12N5/10

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	J. GEN. VIROL., vol. 75, 1994, pages 3547-3553, XP002003353 M.F. BATESON ET AL.;: "Papaya ringspot potyvirus: isolate variability and the origin of PRSV type P (Australia)" see abstract, introduction, Tables 1 and 2 and Figures 1 and 2. ---	1-18
Y	PHYTOPATHOLOGY, vol. 84, 1994, pages 1359-1366, XP002003354 P.F. TENNANT ET AL.;: "Differential protection against papaya ringspot virus isolates in coat protein gene transgenic papaya and classically cross-protected papaya" see abstract and material and methods. --- -/--	1-18

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
20 May 1996	07.06.96

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl, Fax (+ 31-70) 340-3016	Yeats, S

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Internation. pplication No
PCT/US 95/07272

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO,A,90 02184 (THE UPJOHN COMPANY) 8 March 1990 see Examples 2 and 5-8 and claims. ---	1-18
Y	PHYTOPATHOLOGY, vol. 84, 1994, pages 1205-1210, XP002003355 C.H. WANG ET AL.;: "Comparison of the nuclear inclusion b protein and coat protein genes of five papaya ringspot virus strains distinct in geographic origin and pathogenicity" cited in the application see abstract and Figure 1. ---	1-18
Y	US,A,5 162 601 (THE UPJOHN COMPANY) 10 November 1992 see Example 1 and Claims. -----	1-18

1

INTERNATIONAL SEARCH REPORT

International application No
PCT/US 95/07272

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
WO-A-9002184	08-03-90	AU-B-	639891	12-08-93
		AU-B-	3970489	23-03-90
		AU-B-	634168	18-02-93
		AU-B-	3987089	23-03-90
		CA-A-	1332718	25-10-94
		CA-A-	1329561	17-05-94
		DE-D-	68915282	16-06-94
		DE-T-	68915282	29-09-94
		EP-A-	0429478	05-06-91
		EP-A-	0429483	05-06-91
		EP-A-	0693555	24-01-96
		EP-A-	0699757	06-03-96
		JP-T-	4500151	16-01-92
		JP-T-	4500152	16-01-92
		WO-A-	9002189	08-03-90
US-A-5162601	10-11-92	NONE		